g 451
31.8.76

3 Founded in 1887 by G. Stanley Hall 1970

THE AMERICAN
URNAL OF PSYCHOLOGY

EDITED BY

LLOYD G. HUMPHREYS
M. E. BITTERMAN
CHARLES W. ERIKSEN
BENTON J. UNDERWOOD

WITH THE COOPERATION OF

L M. DALLENBACH, WILLIAM EPSTEIN, RAND EVANS, ROBERT FOX,
NDELL GARNER, J. P. GUILFORD, HARRY HELSON, E. R. HILGARD,
IAN HOCHBERG, FRANCIS W. IRWIN, F. NOWELL JONES, JOHN KRAUS-
F, G. L. KREEZER, W. C. H. PRENTICE, ALLAN WAGNER

Vol. 83
1970

PUBLISHED QUARTERLY BY THE UNIVERSITY OF ILLINOIS PRESS
Urbana, Illinois 61801

© 1970 by the Board of Trustees of the University of Illinois
Manufactured in the United States of America

TABLE OF CONTENTS

ARTICLES AND NOTES

Amster, H., Keppel, G., and Meyer, A. Learning and Retention of Letter Pairs as A Function of Association Strength 22
Anderson, N. H. *See* Massaro, D. W.
Attneave, F. *See* Olson, R. K.
Baust, R. F. *See* Lindauer, M. S.
Berger, S. M., Irwin, D. S., and Frommer, G. P. Electromyographic Activity during Observational Learning 86
Bintz, J. A Variable-Interval Generator with Automated Change in Mean Interval 283
Birnbaum, I. M. Recall of a Word List after Interpolated Reading 412
Brislin, R. W., and Leibowitz, H. W. The Effect of Separation between Test and Comparison Objects on Size Constancy at Various Age-levels 372
Chitayat, D. *See* Mack, A.
Christ, R. E. Some Effects of Stimulus-Exposure Time on Choice-Reaction Time 264
Colegate, R. L., and Eriksen, C. W. Implicit Speech as an Encoding Mechanism in Visual Perception 208
Coren, S. *See* Festinger, L.
Culver, C. M. Errors in Tactile Localization 420
Curtis, C. D. *See* Kanak, N. J.
Cvetkovich, G. *See* Rappoport, L.
Dabice, M. *See* Sullivan, R.
Dey, M. K. Generalization of Position Association in Rote Serial Learning 248
Duncan, C. P. The Effect of the Shape of Distribution of Response Strength on Guessing Errors in Problem Solving 535
Epstein, W., and Morgan, C. L. Adaptation to Uniocular Image Magnification: Modification of the Disparity-Depth Relationship 322
Erber, J. T. *See* Kausler, D. H.
Eriksen, C. W. *See* Colegate, R. L.
Eriksen, C. W., and Rohrbaugh, J. W. Some Factors Determining Efficiency of Selective Attention 330
Festinger, L., Coren, S., and Rivers, G. The Effect of Attention on Brightness Contrast and Assimilation 189
Fishbein, H. D., Haygood, R. C., and Frieson, D. Relevant and Irrelevant Saliency in Concept Learning 544
Freund, J. S. *See* Underwood, B. J.
Frieson, D. *See* Fishbein, H. D.
Frommer, G. P. *See* Berger, S. M.

Gardner, G. T. *See* Weintraub, D. J.
Garner, W. R. *See* Preusser, D.
Gogel, W. C. *See* Mershon, D. H.
Gottwald, R. L. *See* Preusser, D.
Hall, J. F., and Mays, L. The Interrelationships among a Number of Association Measures 436
Hall, J. W., and Kozloff, E. E. False Recognitions as a Function of Number of Presentations 272
Haygood, R. C. *See* Fishbein, H. D.
Hintz, J. M., and Nelson, T. M. Golden Section: Reassessment of the Perimetric Hypothesis 126
Holding, D. H. A Line Illusion with Irrelevant Depth Cues 280
Irwin, D. S. *See* Berger, S. M.
Jones, M. B. Rate and Terminal Processes in Skill Acquisition 222
Kanak, N. J., and Curtis, C. D. Unlearning of Context-Item and Specific Associations in the W_1-R_1, W_2-R_2 Paradigm 256
Kappauf, W. E. The Sixty-sixth Annual Meeting of the Society of Experimental Psychologists 630
Katahn, M. *See* Snyder, C. R.
Kausler, D. H., Erber, J. T., and Olson, G. A. Taxonomic Instances as Right or Wrong Items and Selection Strategies in Verbal-Discrimination Learning 428
Kaye, H. *See* Kempler, B.
Kempler, B., and Kaye, H. Agreement among Subjects in Choosing Trigram Labels for Random Shapes 68
Keppel, G. *See* Amster, H.
Kerr, T. P. *See* Thurlow, W. R.
Kozloff, E. E. *See* Hall, J. W.
Krueger, L. E. The Effect of Acoustic Confusability on Visual Search .. 389
Leibowitz, H. W. *See* Brislin, R. W.
Lester, G., and Morant, R. B. Apparent Sound Displacement during Vestibular Stimulation 554
Lindauer, M. S., and Baust, R. F. Instructions and Knowledge of the Situation on Brightness Perception 130
Mack, A., and Chitayat, D. Eye-Dependent and Disparity Adaptation to Opposite Visual-Field Rotations 352
Mackay, S. A. *See* Martin, E.
Martin, E., and Mackay, S. A. A Test of the List-Differentiation Hypothesis 311
Massaro, D. W., and Anderson, N. H. A Test of a Perspective Theory of Geometrical Illusions 567
Masters, L. Knowledge of Results and Item Associability in Paired-Associates Learning 76
Mays, L. *See* Hall, J. F.
McGuigan, F. J. Covert Oral Behavior as a Function of Quality of Handwriting 377
Mershon, D. H., and Gogel, W. C. Effect of Stereoscopic Cues on Perceived Whiteness 55
Mewhort, D. J. K. Guessing and the Order-of-Approximation Effect .. 439

Meyer, A. *See* Amster, H.
Michalewski, H. An Unobtrusive Photoconductive Eyeblink Transducer for Human Subjects 283
Miller, A. L., Pedersen, V. M., and Sheldon, R. J. Magnitude Estimation of Average Length: A Follow-Up 95
Morgan, C. L. *See* Epstein, W.
Morant, R. B. *See* Lester, G.
Mueller, J. H., and Pickering, J. P. Correction Procedures in Three-Alternative Verbal-Discrimination Lists with Intertriad Associations .. 401
Nelson, T. M. *See* Hintz, J. M.
Olson, G. A. *See* Kausler, D. H.
Olson, R. K., and Attneave, F. What Variables Produce Similarity Grouping? 1
Parker, J. F. *See* Postman, L.
Parks, T. E. A Control for Ocular Tracking in the Demonstration of Postretinal Visual Storage 442
Pedersen, V. M. *See* Miller, A. L.
Pickering, J. P. *See* Mueller, J. H.
Postman, L., and Parker, J. F. Maintenance of First-List Associations during Transfer 171
Preusser, D., Garner, W. R., and Gottwald, R. L. Perceptual Organization of Two-Element Temporal Patterns as a Function of their Component One-Element Patterns 151
Rappoport, L., and Cvetkovich, G. Effects of Reward Structure and Cognitive Difference in a Mixed-Motive Two-Person Conflict Situation 119
Rivers, G. *See* Festinger, L.
Rohrbaugh, J. W. *See* Eriksen, C. W.
Sheldon, R. J. *See* Miller, A. L.
Snyder, C. R., and Katahn, M. The Relationship of State Anxiety, Feedback, and Ongoing Self-reported Affect to Performance in Complex Verbal Learning 237
Solso, R. L. Stimulus Selection by Five-, Seven-, and Nine-Year-Old-Children 216
Sperling, G. Binocular Vision: A Physical and a Neural Theory 461
Stafford, J. W. Thomas Verner Moore: 1877–1969 286
Sullivan, R., Warren, R., and Dabice, M. Minimal Aversion Thresholds for White Noise: Adaptation 613
Teghtsoonian, M. *See* Teghtsoonian, R.
Teghtsoonian, R., and Teghtsoonian, M. The Effects of Size and Distance on Magnitude Estimations of Apparent Size 601
Thurlow, W. R., and Kerr, T. P. Effect of a Moving Visual Environment on Localization of Sound 112
Torrey, C. C. Trace Localization and the Recognition of Visual Form .. 591
Turnage, T. W. Temporal Shifts in the Verbal Behavior of Introductory Psychology Students 621
Underwood, B. J., and Freund, J. S. Word Frequency and Short-term Recognition Memory 343
Valle, F. P. Effects of Strain, Sex, and Illumination on Open-Field Behavior of Rats 103

Warren, R. *See* Sullivan, R.
Weintraub, D. J., and Gardner, G. T. Emmert's Laws: Size Constancy vs. Optical Geometry 40
Wenderoth, P. M. A Visual Spatial Aftereffect of Surface Slant 576

BOOK REVIEWS

(The reviewer's name appears in parentheses after the title of the work.)

Ausubel, D. P. *Educational Psychology: A Cognitive View* (H. Cook) .. 303
Bancroft, T. A. *Topics in Intermediate Statistical Methods*. Vol. I. (V. E. McGee) 140
Bauer, R. A., and Gergen, K. J., eds. *The Study of Policy Formation* (K. Weick) 458
Beecroft, R. S. *Classical Conditioning* (D. C. Raskin) 293
Bock, R. D., and Jones, L. V. *The Measurement and Prediction of Judgment and Choice* (D. Krantz) 289
Borden, G. A., Gregg, R. B., and Grove, T. G. *Speech Behavior and Human Interaction* (J. R. Davitz) 454
Brier, R. *See* Rhine, J. B.
Burroughs, W. S. *Junkie* (R. R. Bush) 298
Butcher, H. J. *Human Intelligence: Its Nature and Measurement* (D. B. Harris) 455
Cohen, J. *Information Theory and Esthetic Perception* (P. Vitz) 451
Corso, J. F. *The Experimental Psychology of Sensory Behavior* (E. Carterette) 445
Driver, M. J. *See* Schroder, H. M.
Duncan, H. D. *Symbols and Social Theory* (A. Hershkowitz) 639
Eden, M. *See* Kolers, P. A.
Evans, W. O., and Nathan Kline, eds. *The Psychopharmacology of the Normal Human* (R. R. Bush) 296
Fairweather, G. W. *Methods for Experimental Social Innovation* (R. S. Malpass) 302
Gergen, K. J. *See* Bauer, R. A.
Gilman, A. *See* Goodman, L. S.
Goodman, L. S., and Gilman, A., eds. *The Pharmacological Basis of Therapeutics* (R. R. Bush) 298
Gregg, R. B. *See* Borden, G. A.
Gross, C. G., and Zeigler, H. P. *Readings in Physiological Psychology: Neurophysiology/Sensory Processes* (D. C. Hood) 453
Grove, T. *See* Borden, G. A.
Gurowitz, E. M. *The Molecular Basis of Memory* (C. Cotman) 452
Haber, R. N., ed. *Contemporary Theory and Research in Visual Perception* (J. C. Hay) 138

Hellmuth, J. *Learning Disorders*. Vols. 1 and 2. (K. G. Scott) 142
Hoffer, A., and Osmond, H. *The Hallucinogens* (R. R. Bush) 298
Houston, J. *See* Masters, R. E. L.
Jones, L. V. *See* Bock, R. D.
Kenshalo, D. R., ed. *The Skin Senses: Proceedings of the First International Symposium on the Skin Senses* 145
Kline, N. *See* Evans, W. O.
Köhler, W. *The Task of Gestalt Psychology* (S. Coren) 291
Kolers, P. A., and Eden, M., eds. *Recognizing Patterns: Studies in Living and Automatic Systems* (D. R. Brown) 292
Lazarus, R. S. *Psychological Stress and the Coping Process* (W. W. Lambert) 634
Lingeman, R. R. *Drugs from A to Z: A Dictionary* (R. R. Bush) 298
Lunzer, E. A., and Morris, J. F., eds. *Development in Human Learning* (J. Church) 294
Marx, M. *Learning: Processes* (H. Rachlin) 452
Masters, R. E. L., and Houston, J. *The Varieties of Psychedelic Experience* (R. R. Bush) 299
Medical Economics, Inc. *Physicians' Desk Reference to Pharmaceutical Specialties and Biologicals* (R. R. Bush) 299
Morris, J. F. *See* Lunzer, E. A.
Näätänen, R. *Selective Attention and Evoked Potentials* (L. Kaufman) .. 631
Osmond, H. *See* Hoffer, A.
Plutchik, R. *Foundations of Experimental Research* (B. R. Bugelski) 301
Ravanette, A. T. *Dimensions of Reading Difficulties* (A. Chernoff and J. Williams) 142
Rhine, J. B., and Brier, R. *Parapsychology Today* (D. A. Summers) 304
Rosevear, J. *Pot: A Handbook of Marihuana* (R. R. Bush) 298
Schlesinger, I. M. *Sentence Structure and the Reading Process* (J. W. Torry) 637
Schroder, H. M., Driver, M. J., and Streufert, S. *Human Information Processing: Individuals and Groups Functioning in Complex Situations* (R. Hyman) 136
Schuster, C. R. *See* Thompson, T.
Sem-Jacobsen, C. W. *Depth-Electrographic Stimulation of the Human Brain and Behavior* (W. A. Wilson, Jr.) 143
Simon, H. A. *The Sciences of the Artificial* (E. Galanter) 139
Soloman, D. *The Marihuana Papers* (R. R. Bush) 298
Smith, H. B. *Social Psychology and Human Values* (H. Alker) 457
Streufert, S. *See* Schroder, H. M.
Szasz, T. S. *Law, Liberty, and Psychiatry* (R. Singer) 141
Thompson, T., and Schuster, C. R. *Behavioral Pharmacology* (C. Kornetsky) 640
Warren, R. M., and Warren, R. P. *Helmholtz on Perception: Its Physiology and Development* (J. Beck) 448
Warren, R. P. *See* Warren R. M.
Welford, A. T. *Fundamentals of Skill* (J. A. Adams) 299
Zeigler, H. P. *See* Gross, C. G.

Diary No. 152
Date 6/11/70
File No. Library
Bureau Ednl. Psy. Research

WHAT VARIABLES PRODUCE SIMILARITY GROUPING?

Richard K. Olson and Fred Attneave, *University of Oregon*

Abstract. The effectiveness of several configurational or textural variables in producing similarity grouping was investigated by measuring the time it took *S*s to locate the disparate quadrant in a circular array. The 'figure' and 'ground' quadrants differed in (1) slope of elements, (2) linearity of elements, or (3) the way in which the same slopes were combined into angles. Differences in slope produced excellent grouping; differences in linearity were less effective; and differences in arrangement of slopes were generally quite ineffective. Grouping was also dependent on orientation of the whole array: horizontals and verticals gave better grouping than diagonals, and even arrangement of slopes gave moderately effective grouping when angles with horizontal and vertical axes of symmetry were opposed to each other. Conditions of head tilt were used to determine whether the grouping effects were dependent on retinal or gravitational orientation. For the angles, only retinal orientation seemed important; but for line segments and paired dots varying in slope, both retinal and gravitational orientation had significant effects. The results disconfirm the hypothesis that grouping occurs entirely on a projection level, independently of constancy mechanisms. It is suggested that grouping and segregation depend on identities and differences between descriptors which in turn represent relationships between the stimulus array and an internal Cartesian reference system.

Similarity is one of the most important of Wertheimer's principles of perceptual grouping, but only recently, with the work of Julesz and Beck, has any appreciable curiosity been displayed about the kinds of similarity that are effective in this respect.[1] It should, of course, be understood that grouping always implies the complementary process of segregation. Grouped items must be at once similar to one another and different from extraneous items; in analysis-of-variance terminology, within-group variance must be

Received for publication September 22, 1969. The research was supported in part by the Air Force Office of Scientific Research (Grant 973-66) and in part by the Advanced Research Projects Agency of the Department of Defense (Contract F44620-67C-0099).

[1] M. Wertheimer, Untersuchungen zur Lehre von der Gestalt, *Psychol. Forsch.*, 4, 1923, 301–350; B. Julesz, Texture and visual perception, *Sci. Amer.*, 212, 1965, 38–48; J. Beck, Perceptual grouping produced by line figures, *Percept. & Psychophys.*, 2, 1967, 491–495.

low relative to between-group variance on one or more stimulus variables. But *what* stimulus variables may thus produce similarity grouping?

There is no doubt that brightness functions as a grouping variable; likewise, the grouping effect of hue is demonstrated in pseudoisochromatic tests of color vision. Where we begin to encounter real difficulties is in attempting to assess the role of configurational variables. There is a widespread acceptance of the notion that 'shape' gives grouping, based on such familiar examples as

xxooxxoo ,

and on elaborations thereof in which a figure area is filled with one repeated letter and a ground area with another. However, examples of this type neglect even the most rudimentary controls for the possible effects of lower-order variables. Since no two different letters are likely to contain the same amount of ink, average brightness is confounded with shape. Moreover, if (as we shall presently see) line slope is an effective grouping variable, it is quite possible that letters group or segregate because they contain similar or different distributions of slopes and that the arrangement or configuration of component lines is irrelevant.

In some situations the inherent difficulty of adequate control is very great. For illustrative purposes, let us consider a particularly formidable example. Grouping can be produced by something that may loosely be called texture grain, which on close inspection turns out to involve a whole cluster of variables. Suppose, for example, that a matrix is divided into two regions: in the first, cells are filled randomly with some given probability; in the second, they are filled with the same probability (so that average brightness is controlled) but with further constraints such that cells of the same state tend to form sizeable clumps. Phenomenal grouping and segregation result, and the observer may readily tell us that the first region has a finer texture than the second; but neither this description nor the method of generation enables us to specify precisely the psychological variable or variables on which the grouping depends. One possibility is that size or extent (of like-colored areas) is the important variable. Another hypothesis—a quite plausible one in view of what we now know about lateral inhibition—is that amount of contour is a basis for grouping. (Clumping reduces amount of contour.)

Next, let us set up a display (Fig. 1) in which amount of contour

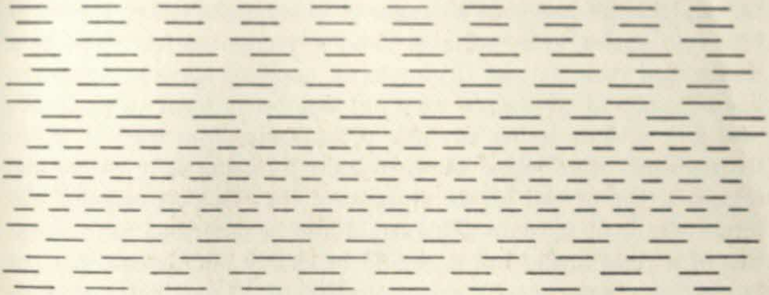


FIG. 1. Grouping based on length with amount of contour and average brightness held constant

is approximately constant, as is average brightness, but extent (length of line segments) varies from one region to another. It may appear that the good grouping which results can be attributed only to length, but unfortunately the situation is more complicated. In addition to (*a*) element length, some other variables are (*b*) distances between elements, *i.e.* average size of compact white areas; (*c*) object density, *i.e.* number of segments per unit area; (*d*) subjective brightness as a result of nonlinear brightness averaging (since what is held constant is the purely physical variable of amount of ink per unit area); (*e*) subjective amount of contour as a result of nonlinear averaging; and (*f*) variability in brightness or in amount of contour from one summation region to another (since if either brightness or amount of contour is summated or averaged within fairly small regions, these averages will be more homogeneous where line segments are short than where they are long). Note well the control problem inherent in such an interlocking set of variables. If, for example, one holds object density constant (and still varies length), both brightness and amount of contour become uncontrolled, and so on. No experimental strategy to untie these variables is evident to the authors, and we may have to settle for the weak conclusion that grouping is produced by *some* variables within the cluster.

The studies here reported investigate another set of configurational variables which, though more amenable to experimental analysis than those of texture density, still necessitate unusual attention to problems of control. These variables were (1) *slope, i.e.* straight line segments of one orientation or tilt as opposed to straight line segments of another orientation or tilt; (2) *linearity,* *i.e.* straight line segments as opposed to segments either curved or bent into angles, every segment randomly oriented lest slope be a confounded variable; and (3) *angle, i.e.* opposing ways in which two line segments of given slope were put together to form an angle. We used four such angles for any pair of segments—four ways in which the segments were joined at their ends. By using only orthogonal segments, the overall extent or span of the resulting angular elements was kept constant. This could also be described as 'orientation of a right angle,' but it should be clearly understood that our interest was in the possible grouping effect of the particular *arrangements,* or *configurations,* of the slopes in these angles. Fig. 2 shows some examples of arrays based on these variables. In every case one quadrant of a circle is occupied by elements that differ, on the variable under consideration, from the elements in the remainder of the circle. The observer's task was to locate as quickly as possible the disparate quadrant, which could take any one of *eight* positions (compare *1* and *13* in Fig. 2). The rationale for using latency as a response measure was simply that good grouping seems to occur 'immediately,' whereas a search process obviously requires time.

EXPERIMENT I

Method

Subjects. The eight *S*s, four men and four women, were paid undergraduate volunteers. All met a criterion of 2/2 acuity without correction, and none showed astigmatism on the Lancaster-Regan test.

Materials. The 14 classes of element combinations used in the experiment are shown in Table I. The left element of each pair in the table was the 'figure' element, repeated in one quadrant of an 80-element matrix. The right element of the pair was the 'ground' element, repeated in the remaining three quarters of the matrix (see examples in Fig. 2).

A quasi-random method of positioning stimulus elements was designed to provide an approximately uniform density or spacing of elements and at the same time to avoid neat row-and-column arrangements that would introduce collinearity and border effects as confusing variables. Within each of 80 square cells, arranged in a matrix with an approximately circular border, one element was placed, but its location within that cell was determined randomly, with contraints such that it could never quite touch any of its neighbors. Under these rules an element was free to vary, both vertically and horizontally, by a little less than twice its linear extent.

In the original india-ink drawings, all elements were contained in a circle of 7.5 in. diameter. In Categories 1-4, all the line segments were .25 in. long;

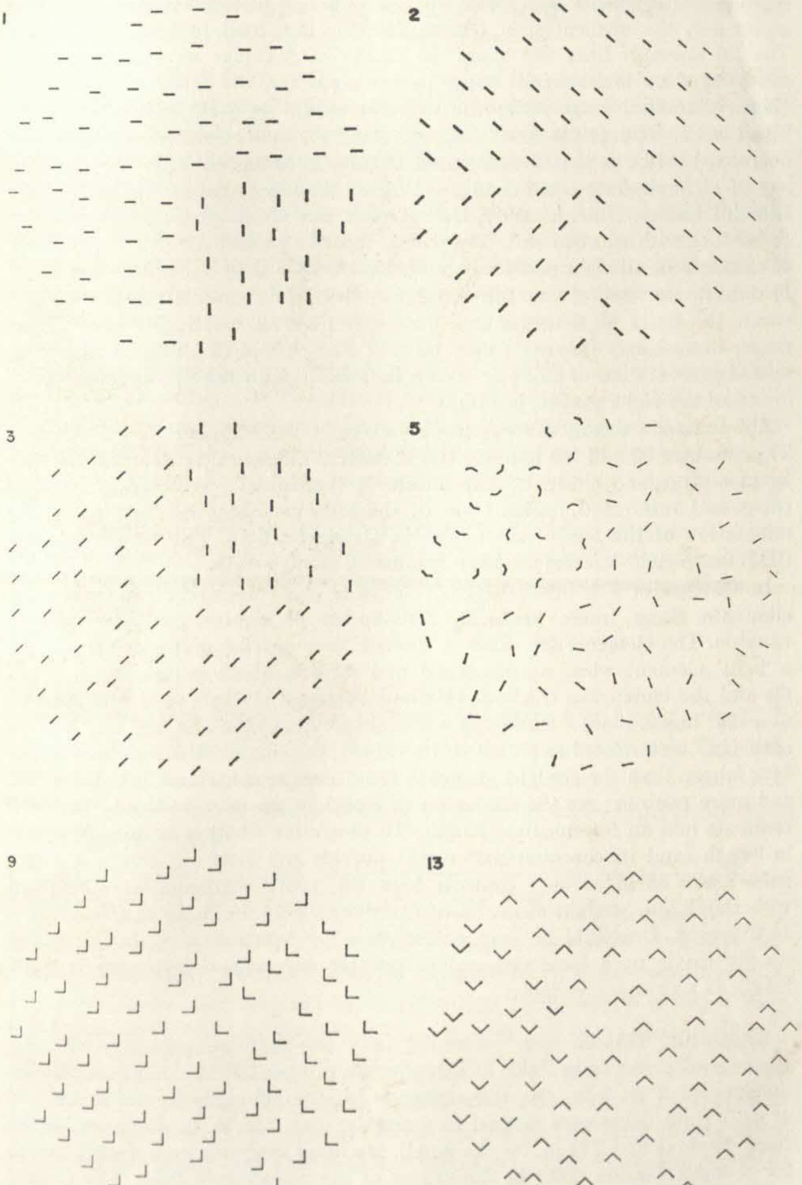


Fig. 2. Sample displays from Experiment I. Numbers refer to corresponding categories in Table I. Other categories are illustrated in a 45° rotation of the figure; note especially that Category 9 rotates into Category 12.

so were the segments that made up the angles in Categories 9–14. In Categories 5–7, the element span (linear distance from end to end) was .25 in. The thickness of lines was about .03 in. These drawings were photographed and printed on high-contrast paper in two sizes: one the same as the original (7½-in. diameter, 'large set'); the other one-third as great (2⅝-in. diameter, 'small set'). The prints were aligned, centered, and mounted on octagonal fiberboard cards, so that the disparate (figure) quadrant could be presented in any of eight discrete positions depending on the placement of the card in the stimulus holder. Note, however, that a given card could go from one category to another with rotation and, conversely, that the presentation of a given pair of elements in all eight positions could require more than one card. (See Table I; only in the case of Categories 5–8 was there a one-to-one relationship between the pairs of elements and the cards over all rotations). Sparing the reader unnecessary details, it may be said that the set of stimulus cards permitted presentation of any pair shown in Table I, with the disparate quadrant in any of the eight possible positions.

Abbreviated category descriptions are given in Table I, next to the elements. The numbers 90 and 180 indicate the rotational difference, in degrees, between figure and ground elements. The letters H (horizontal), V (vertical), and D (diagonal) refer to the orientation of the lines in Categories 1–4 and to the orientation of the legs of the angles in Categories 9–14. The letters RL and UD refer to right-left and up-down symmetry of angle pairs.

In Categories 5–8, orientation of elements was completely randomized to eliminate slope (more precisely, distribution of slopes) as a confounded variable. The elements described as 'curved' were quarter circles. A curved and a 'bent' element when superimposed had three common points, at the middle and the ends; thus the bent elements consisted of short segments meeting at a 135° angle. Since a curve and a straight element were also equated in total span (*i.e.* were related as an arc to its chord), the curved elements were about 14% *longer* than the straight elements (and likewise contained 14% more ink, and more contour; see the discussion of length in the introduction). The bent elements had an intermediate length. To determine whether such a difference in length (and its concomitants) might provide any basis for grouping, Category 8 was included as a control: here the 'short' segments were identical with the .25-in. straight elements of Categories 5–7; the 'long' segments were 14% longer. It should be emphasized that the function of Category 8 was merely to tie up a loose end, not to provide any general information about length as a grouping variable.

Apparatus. Stimuli were presented in a two-field tachistoscope with exposure and preexposure fields at a luminance of about 7 mL. and at a viewing distance of 2 ft. 9 in. (At this distance an inch subtends a visual angle of 1° 44′.) Both fields were masked to a circular area, 7.75 in. in diameter for the large displays and 2.75 in. for the small. No other contours were visible except for a small fixation point in the center of the preexposure field. The display holder allowed the octagonal stimulus cards to be inserted in any of eight stable positions. A headrest kept *S*'s head upright and permitted viewing with the right eye only. Both the onset and the duration of stimulus exposure were controlled by *S*: when he pressed a switch with his hand, the stimulus field and a clock were turned on; when he released the switch, the clock stopped and the field reverted to the preexposure state.

Procedure. The *S* was told that the purpose of the experiment was to study the grouping properties of certain stimuli. He was shown sample displays of each category (outside the apparatus), and the quadrant organization with the eight possible positions was explained. He was instructed to fixate the preexposure dot before each trial, to press the switch when ready, and to release it the moment he could identify with certainty the location of the disparate quadrant. (Guessing was discouraged, and errors were extremely rare.) After releasing the switch, he indicated verbally the position of the quadrant. He was then asked to rate the strength of grouping on a seven-point scale, with 1 indicating 'very clear grouping' and 7 'very poor grouping.'

Each *S* was run about an hour a day for nine days. The first day was considered a practice session, and its data were not used. On every subsequent session, *S* was presented with two stimuli from each category, one large and one small. The two sizes were presented in separate series within a session, with order counterbalanced over sessions. Selection of the particular pair within a category was randomized over sessions with the limitation that each pair be used equally often. (It is obvious from Table I that some categories contain more alternative pairs than others.) Position of the figure quadrant was also randomly permuted. If *S* did not locate the disparate quadrant within 50 sec., the trial was terminated and the time was arbitrarily recorded as 50 sec.

Results

Mean reaction times (RT) and ratings were obtained for each *S* by categories. Means over *S*s for each category are given in Table I. Table II shows the results of 17 *t*-test comparisons of theoretical interest, their means, descriptions, *t* values ($df = 7$), and confidence levels (two-tailed).

Ratings were analyzed statistically only with respect to their correlation with RT, for which $r = .925$. The ratings added little information to the more objective measure, nor have we any assurance that they were not directly influenced by time requirements. It seems likely that *S*s paid little attention to category labels ('very good grouping,' 'poor grouping,' etc.) and simply accommodated the scale to the range of stimuli as raters typically do. The size of the stimulus display had no demonstrable effect on RTs for lines (Categories 1–4). For angles (Categories 9–14), responses were moderately faster on the smaller displays. We have no way of knowing, without further experimentation, whether this difference was due to overall display size per se or to the size or spacing of component elements. On the straight-curved-bent pairings (Categories 5–7), RTs to the small displays were so long as to suggest

TABLE I

MEAN REACTION TIMES (RT) AND RATINGS FOR DISPLAYED ELEMENT PAIRS: EXPERIMENT I

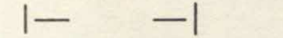
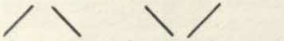
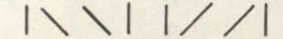
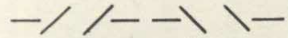
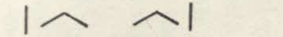
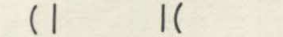
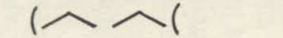
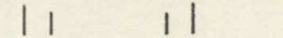
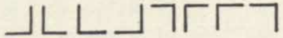
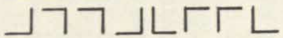
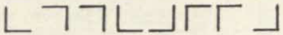
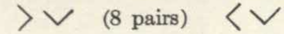
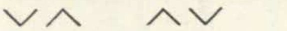
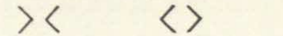
Category	Element pairs	Description	Large set		Small set		Hypothesized descriptor
			RT	Rating	RT	Rating	
1		HV 90	.86	1.06	.93	1.00	H vs. V
2		D 90	1.42	1.14	1.29	1.18	Not H not V vs. Not H not V
3		VD	1.05	1.05	1.18	1.05	V vs. Not V
4		HD	1.05	1.03	1.01	1.01	H vs. Not H
5		Straight-Bent	2.52	2.46	7.56	3.65	Linear vs. Nonlinear (strong)
6		Curved-Straight	3.62	2.54	9.69	3.78	Linear vs. Nonlinear (weak)
7		Curved-Bent	7.32	3.64	41.21	6.32	Nonlinear (weak) vs. Nonlinear (strong)
8		Long-Short	26.11	5.36	32.35	5.78	
9		HV 90, RL	4.09	2.36	3.59	2.12	HV vs. HV
10		HV 90, UD	3.69	2.26	2.78	2.12	HV vs. HV
11		HV 180	3.83	2.31	3.65	2.26	HV vs. HV
12	 (8 pairs)	D 90	2.69	1.85	1.94	1.70	H axis vs. V axis
13		D 180, UD	4.26	2.44	3.96	2.26	V axis vs. V axis
14		D 180, RL	4.16	2.33	3.27	2.07	H axis vs. H axis

TABLE II

STATISTICAL COMPARISON: SELECTED CATEGORIES, EXPERIMENT I

Description of comparison	Respective means		Diff.	t
Large vs. Small, all lines (1–4)	1.09	1.10	.01	.08
Large vs. Small, all angles (9–14)	3.79	3.20	.59	2.62*
All angles (9–14) vs. All lines (1–4)	3.49	1.10	2.39	11.41**
Straight-bent (5) vs. All lines (1–4)	2.52	1.10	1.42	7.53**
All angles (9–14) vs. Straight-bent (5)	3.49	2.52	.97	3.93**
Straight-curved (6) vs. Straight-bent (5)	3.62	2.52	1.10	4.58**
Straight-curved (6) vs. All lines (1–4)	3.62	1.10	2.52	7.20**
Straight-curved (6) vs. All angles (9–14)	3.62	3.49	.13	.44
Line pairs				
D 90 (2) vs. HV 90 (1)	1.35	.89	.46	3.48*
VD & HD 45 (3–4) vs. HV 90 (1)	1.07	.89	.18	1.75
D 90 (2) vs. VD & HD 45 (3–4)	1.35	1.07	.28	1.75
Angle pairs				
HV 90 (9–10) vs. HV 180 (11)	3.53	3.74	.21	.26
HV 90 RL (9) vs. HV 90 UD (10)	3.84	3.23	.67	1.97
D 180 UD (13) vs. D 180 RL (14)	4.11	3.72	.39	.99
D 180 (13–14) vs. HV 180 (11)	3.91	3.74	.17	.47
D 180 (13–14) vs. D 90 (12)	3.91	2.31	1.60	5.57**
HV 90 (9–10) vs. D 90 (12)	3.53	2.31	1.22	3.38**

Note: Categories shown in parentheses.
* $p < .05$.
** $p < .01$.

strongly the pressing of acuity limitations, in which we are not immediately interested. For this reason, all comparisons involving these elements were based on data from the large displays only.

The most obvious gross differences in the results were between the three major bases on which the elements could be grouped. Contrasting slopes (the straight lines in Categories 1–4) produced excellent grouping, with RTs of the order of 1 sec.[2] At the other extreme, responses to contrasting arrangements of the same slopes (the angles in Categories 9–14) typically required 3 or 4 sec.; location of the disparate quadrant was in no sense 'immediate' but rather the outcome of a search task. The best contrast between linear and nonlinear elements (the straight-bent elements in Category 5) was an intermediate case, requiring about 2.5 sec. The reader is invited to take another look at Fig. 2 and compare the *S*s' results with his own phenomenal impressions. In the study already cited (see n. 1), Beck, using rather different methods, had previously concluded that slope is the major variable in grouping based upon 'shape.' Our results support this conclusion.

---

[2] It can be shown otherwise that most of this is 'response' time: slope still gives good grouping at exposure times of .1 sec. or less.

Let us now consider differences within these broad classes, looking first at Categories 1–4. The horizontal-vertical pairs in Category 1 gave significantly faster grouping than the Category 2 pairs of diagonals (which also differed by 90°); the Category 3 and 4 pairs of a diagonal with either a horizontal or a vertical (differing by only 45°) were in between. Although significance can not be demonstrated for the differences between these intermediate cases and the extremes, the ordering is quite a coherent one, indicating that absolute values of orientation can be more important to grouping than differences in orientation.

The comparisons among Categories 9–14 show simply that there was something unique about Category 12, for which the mean RT was only 2.31 sec. This is perhaps the only one of the pairs of angles for which a casual observer might be inclined to claim some 'immediacy' of grouping (see Fig. 2). To account for this unique case, let us first examine the possibility that slope—which we know produces good grouping—was inadequately controlled. Ostensibly, it was controlled by employing exactly the same (diagonal) slopes in each member of the pair and by varying only the way in which they were joined. However, it may be argued that an element with vertex at top or bottom has an overall horizontal orientation, whereas one with vertex at left or right has an overall vertical orientation. In fact, the former has its greatest extent in the horizontal, the latter in the vertical. Likewise, in terms of slope analyzers of the type described by Hubel and Wiesel, one element might be expected to fit best a vertically oriented (elliptical) receptive field, the other a horizontally oriented field.[3]

By this line of reasoning, elements differing by a 90° rotation (Categories 9, 10, and 12) vary in overall slope as well as in arrangement of slopes and should therefore give better grouping than elements that differ by 180° and in which overall slope is constant (Categories 11, 13 and 14). As noted, however, Category 12 was unique: Categories 9 and 10 (characterized by horizontal and vertical legs; see Table I) were no better grouped than the 180° cases. To retain an explanation in terms of slope analyzers, then, we would have to suppose that many more receptive fields in the human visual system are oriented horizontally and vertically than are oriented diagonally. This possibility cannot be totally rejected, but the weight of previous behavioral evidence seems somewhat against it;[4] furthermore, no such paucity of diagonal analyzers has been evident in the primates that Hubel and Wiesel have studied.[5]

---

[3] D. H. Hubel and T. N. Wiesel, Receptive fields and functional architecture in two non-striate visual areas (18 and 19) of the cat, *J. Neurophysiol.*, 28, 1965, 229–289; D. H. Hubel and T. N. Wiesel, Receptive fields and functional architecture of monkey striate cortex, *J. Physiol.*, 195, 1968, 215–243.

There is some reason, therefore, to seek an explanation at a higher level of the perceptual system. Attneave has recently presented evidence that the visual system describes figures in terms of an essentially Cartesian reference system and that the appearance of an object depends upon its relationship to the descriptive axes.[6] Note that the elements of Categories 9–11 seem phenomenally different, in quite a fundamental way, from those of Categories 12–14.[7] The assumption of a Cartesian frame of reference immediately clarifies this difference. In Categories 9–11, the legs of the angles coincide with Cartesian axes, and it is accordingly very obvious that they are right angles. In Categories 12–14, the orthogonality of the elements almost goes unnoticed, whereas the symmetry of elements, which was quite unobtrusive in Categories 9–11, becomes highly salient.

Suppose that grouping and segregation are based upon certain *descriptors*, some of which represent relationships of elements to a Cartesian reference system, and that descriptors of the simplest such relationships are the ones most easily and quickly abstracted. The hypothesized descriptors listed in the last column of Table I represent particularly rudimentary relationships between the respective elements and the postulated descriptive system. The unique status of Category 12 is attributable to the fact that it contrasts vertical with horizontal axes of symmetry.[8] In no other case of angle pairs is the hypothesized (basic or initial) descriptor different from one member of the pair to the other: in Categories 9–11, all elements are characterized simply as horizontal-vertical combinations, *i.e.* as right angles; and axis of symmetry provides no basis for segregation in either Category 13 or 14.

---

[4] F. Attneave and R. Olson, Discriminability of stimuli varying in physical and retinal orientation, *J. exp. Psychol.*, 74, 1967, 149–157; F. Attneave and K. Reid, Voluntary control of frame of reference and slope equivalence under head rotation, *J. exp. Psychol.*, 78, 1968, 153–159.

[5] See the second paper cited in n. 3 above.

[6] F. Attneave, Triangles as ambiguous figures, this JOURNAL, 81, 1968, 447–453.

[7] See the discussion of the phenomenal distinction between a square and a diamond in E. Mach, *The Analysis of Sensations*, 1959, 106.

[8] Postulation of a descriptor for symmetry about the horizontal is somewhat inconsistent with the evidence (see Attneave, n. 6 above; I. Rock and R. Leaman, An experimental analysis of visual symmetry, *Acta Psychol.*, 21, 1963,

The hypothesis that elements in Categories 12–14 are classified by axis of symmetry is consistent with certain of Attneave's observations on ambiguous triangles (see n. 6). Alternatively, or additionally, however, we might postulate descriptors dependent on coincidence of the greatest extent of the element (or of the line determined by the free ends of the element) with either a vertical or a horizontal reference axis. 'Overall slope,' thus reinterpreted, is perfectly correlated with axis of symmetry and offers a similar basis for the superiority of Category 12. (It has been repeatedly suggested to the authors that combinations of diagonal segments are more readily seen as familiar objects—as 'birds,' 'rooftops,' 'arrows,' and so forth—than are combinations of horizontals and verticals. This observation may be true enough, but it does not contribute to any coherent explanation of the results. If descriptive processes of this type make for good grouping, Category 13—'birds' vs. 'rooftops'—should have been particularly easy, but it was not.)

Returning to single lines (Categories 1–4), we see that the ordering of RTs was likewise consistent with the assumption of a Cartesian descriptive system. If the most basic descriptors are simply 'vertical' and 'horizontal,' there are two such descriptors making for segregation in Category 1; whereas in Categories 3 and 4, there is only one; and in Category 2, none. (Rather obviously, we must suppose that there is some weaker or later descriptor which *does* differentiate between the diagonals and that it in turn is more salient than any of the descriptors of angles.) An explanation of these results in terms of receptive fields would again, as in the case of the angles, require that we assume a preponderance of vertical and horizontal over diagonal analyzers.

171–183) that symmetry is ineffective except about a phenomenal or figural vertical. Actually, it makes no difference to the explanation above to suppose that elements symmetrical about the horizontal are described as asymmetrical. Alternatively, we may suppose that the contrast (in Category 12) is between two figural verticals determined by axes of symmetry, one coincident with the field vertical, the other with the field horizontal. This is a more complicated, but perhaps more accurate, way of characterizing the descriptors suggested above.

The possibility that difficulty is increased by a mirror-image relationship *between* paired elements was also examined. The relative difficulty of paired diagonals (Category 2) might be taken as support for this hypothesis. The difference between Categories 9 and 10 was in the predicted direction, but not to a significant degree [t (7) = 1.97]. Between Categories 13 and 14, the difference was slightly in the wrong direction. This variable would seem, therefore, to have little or no importance.

Within the limitations of this experiment, linearity vs. nonlinearity of elements was at best (in Category 5) a moderately effective grouping variable. The ordering of RTs for Categories 5–7 suggests that straight, curved, and bent segments are related as points on a single continuum. Curved elements are, in effect, weakly nonlinear; bent or angular elements are more strongly so. It is extremely likely that sharper angles would increase the effectiveness of this variable; we chose such obtuse angles, as the reader will recall, to minimize the possible confounding effects of span or length. The unimportance of the length variable confounded in the present case is demonstrated by the very great difficulty of Category 8, in which a difference in length equivalent to that of Category 5 was the only differentiating variable.

EXPERIMENT II

Experiment I showed, in the cases of both lines and angles, that grouping is affected by the orientation of the stimulus field relative to the horizontal and vertical. The theoretical interpretation of these effects might vary considerably depending on whether the important independent variable is retinal orientation or physical (*i.e.* gravitational) orientation. In Experiments II–IV, the two reference systems for orientation were dissociated by having *S*s view the stimuli with their heads tilted 45°.

Method

Subjects. Sixteen paid undergraduate volunteers were run in two groups of eight (four men and four women in each group). All *S*s were screened for visual defects as in Experiment I.

Materials. The materials were the angles that comprised Categories 9–14 in Experiment I.

Apparatus. For Group U, the apparatus used was the same as that in Experiment I. For Group T, the headrest was tilted 45° clockwise, a bite board was added to hold *S*'s head at that tilt, and two small white dots were placed at the gravitational top and bottom of the circular masks.

Procedure. In general, the procedure was the same as in Experiment I except that two matrices of *each* stimulus pair in Categories 9–14, one large and one small, were shown in a session. The *S*s of Group U viewed the stimuli with heads upright, as before. Those of Group T had their heads tilted 45°. They were told to think of the gravitational top of the field as 'up,' referring to the marking dots if necessary. They were allowed to sit back from the apparatus between trials but were required to have their heads in the correct viewing position for at least 20 sec. before each trial, so that effects of torsional eye movements, which are mostly transitory, would be minimal.[9]

Results

For purposes of analysis the six categories (9–14 of Table I) were reduced to four by combining Categories 9 with 10 and 13 with 14, as shown in Fig. 3. The mean RTs graphed in this figure are averages over sizes, *S*s, and quadrants.

The effect of stimulus size (not shown) was similar to that found in Experiment I. Again, ratings were highly correlated with RTs and were not further analyzed. The generally faster RTs in Experiment II are probably attributable to the fact that angles were the only kinds of stimuli presented.

The results (Fig. 3; Table III) can be summarized quite simply. Group U showed exactly the same pattern noted in Experiment I

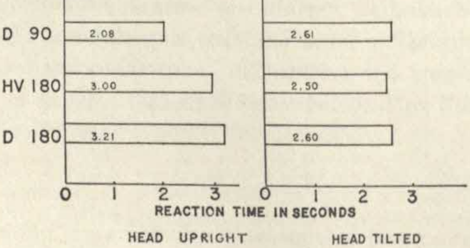


FIG. 3. Reaction times for angles in Experiment II, averaged over subjects

---

[9] K. N. Ogle, Spatial localization according to direction, in H. Douson (ed.), *The Eye*, 3, 1962, 237.

TABLE III

STATISTICAL COMPARISON: EXPERIMENT II

Description of comparison	Respective means		Diff.	*t*
Head upright, Group U				
HV 90 (9–10) vs. HV 180 (11)	2.89	2.99	−.10	.89
D 180 (13–14) vs. D 90 (12)	3.21	2.08	1.13	8.01**
D 180 (13–14) vs. HV 180 (11)	3.21	2.99	.22	.68
HV 90 (9–10) vs. D 90 (12)	2.89	2.08	.81	4.38**
Head tilted, Group T				
HV 90 (9–10) vs. HV 180 (11)	2.10	2.50	−.40	3.10*
D 180 (13–14) vs. D 90 (12)	2.60	2.61	−.01	.09
D 180 (13–14) vs. HV 180 (11)	2.60	2.50	.10	.94
HV 90 (9–10) vs. D 90 (12)	2.10	2.61	−.51	5.00**

Note: Categories shown in parentheses.
* $p < .05$.
** $p < .01$.

(of which it was a partial replication): Category 12 (D 90) was uniquely easier than the others. Group T also showed uniquely faster performance on the combined Categories 9 and 10 (HV 90), which are, however, *retinally* equivalent to Category 12. In short, what mattered here was the retinal rather than the gravitational orientation of the stimulus, despite instructions biased toward the gravitational. The smaller difference in Group T than in Group U might be taken to suggest a combination of proximal and distal effects, but this interaction was not significant. These results are discussed further below.

EXPERIMENT III

Experiment I showed that horizontal vs. vertical lines give better grouping than diagonal vs. diagonal; this difference was further investigated in Experiment III, under both head-upright and head-tilt conditions, and with somewhat different stimuli. Here, pairs of dots, the pairs differing in orientation (see Fig. 4), were used instead of line segments for several reasons: (*a*) simply to extend the generality of the earlier findings; (*b*) to weaken somewhat the basis for grouping and thereby possibly to enhance the

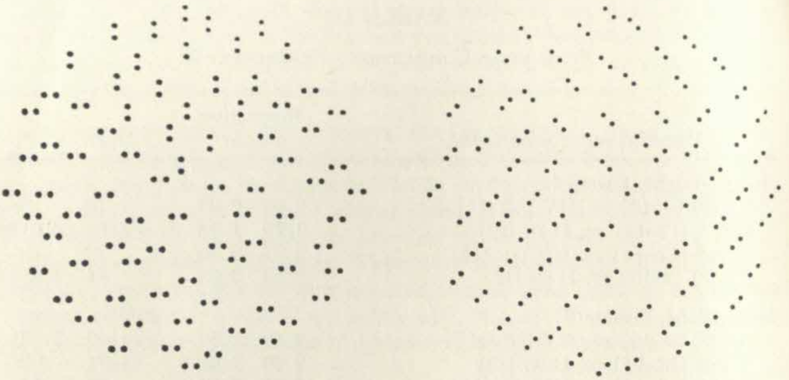


FIG. 4. Sample displays of dot matrices in Experiment III. Separation of paired dots in samples was .085 in.

difference found earlier; (*c*) to investigate the effect of variation in the separation of the paired dots.

Method

Subjects. The *S*s were five men and one woman, from the same population and meeting the same visual requirements as *S*s in Experiments I and II.

Materials. Examples of the new stimulus matrices are shown in Fig. 4. The general plan of Experiment I (one quadrant different from the other three) was again employed. The elements were pairs of dots .033 in. in diameter drawn in india ink on white paper. Eighty such pairs were semirandomly (see Experiment I) positioned in a 2⅝-in.-diameter matrix. Pairs in the 'figure' quadrant were always perpendicular in orientation to those in the 'ground' area. Four different separations of the paired dots, from .085 in. to .108 in. center-to-center (*i.e.* from .052 in. to .075 in. edge-to-edge), were used. For each separation, four matrices were constructed, one for each of the four possible orientations of the pairs in the figure quadrant; by rotation it was thus possible to present any of the four orientations in any of *eight* possible positions.

Apparatus. Apparatus was the same as that in Experiment II.

Procedure. All six *S*s were treated alike except that stimulus order was randomly permuted for each *S* separately. They viewed the stimuli with heads upright (Condition U) and heads tilted 45° (Condition T) alternately over blocks of 21 or 22 trials. In each of two 1-hr. sessions, *S* was presented with all 16 matrices in all eight rotations for a total of 128 trials. Stimulus order was identical for both sessions, but order of head positions was reversed, so that each stimulus was viewed under both conditions. During the T blocks, *S* was allowed to sit back from the apparatus between trials but was requested to keep his head tilted 45°. Procedure was otherwise like that of Experiment II.

Results

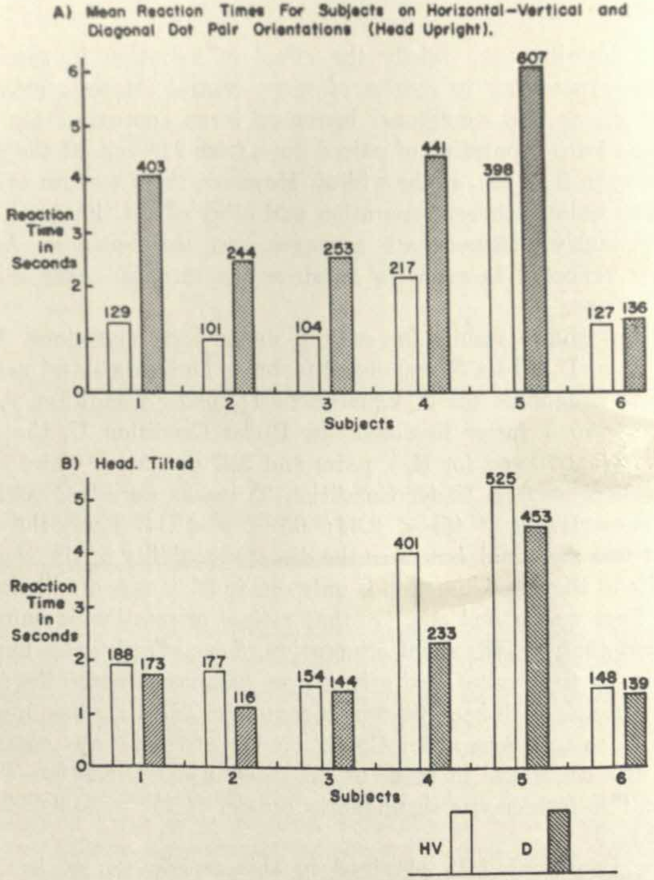
To describe very briefly the effect of variation in separation before proceeding to results of more central interest, mean RT (over all *S*s and conditions) increased in an approximately linear manner with separation of paired dots, from 2.06 sec. at the closest spacing to 3.13 sec. at the widest. However, there was no evidence of interaction between separation and other effects; hence the RTs subsequently discussed are averages over this variable. All the effects reported below were consistent in direction over all four separations.

Fig. 5 shows results for each *S* under both conditions. Under Condition U, all six *S*s responded faster to horizontals and verticals than to diagonals (as in Experiment I); under Condition T, they all responded faster to diagonals. Under Condition U, the mean RTs were 1.79 sec. for H-V pairs and 3.47 sec. for D pairs [t (5) $= 4.47$, $p < .01$]. Under Condition T, means were 2.65 and 2.09 sec. respectively [t (5) < 2.24; $.05 < p < .1$]. Thus, the latter effect was marginal, but since the direct probability of six *S*s giving results in the same direction is only one in 32, it appears likely that the effect was a real one, *i.e.* that retinal orientation is more important than gravitational orientation. *Some* effectiveness must be attributed to gravitational orientation, however, because the reversal of results with head tilt was incomplete. The difference in favor of H-V orientations under Condition U (1.68 sec.) was compared with the difference in favor of diagonals under Condition T (.56 sec.). The former was significantly greater [t (5) $= 3.09$; $.02 < p < .05$].

The very long RTs obtained in this experiment, at least for some *S*s and some stimuli, made us question whether the results could be considered relevant to slope as a *grouping* variable. Although we have no firm criterion for differentiating between grouping and search tasks, RTs of the order of 4–6 sec. (see Fig. 5) certainly seem indicative of the latter. For this reason, we decided to add Experiment IV, which used the line segments of Experiment I, Categories 1 and 2, with a head-tilt as well as a head-upright condition.

EXPERIMENT IV

The basic aim of this experiment, as of the two preceding it, was to dissociate effects of physical and retinal orientation. Additional


A) Mean Reaction Times For Subjects on Horizontal-Vertical and Diagonal Dot Pair Orientations (Head Upright).

B) Head. Tilted

FIG. 5. Reaction times for paired-dot orientations in Experiment III, under two conditions of head tilt

considerations led to a four-factor design including these variables: (1) orientation of elements, whether H vs. V or D vs. D (Table I, Categories 1 and 2); (2) head position, upright or tilted 45°; (3) matrix size, large or small, as in Experiment I; and (4) relationship between elements and quadrants.

The last variable requires some explanation. If the reader refers back to Fig. 2, he will note that the arrays numbered *1* and *2* differ in a respect that is invariant over rotation of either or both. In *1* the elements are either parallel or perpendicular to the figure-ground boundaries, whereas in *2* the elements and boundaries are in a 45° relationship. Belatedly, it appeared to us that the arrangement in *1* made for better segregation of the disparate quadrant than did the arrangement in *2*; therefore this variable, which had previously been buried in the tabulation of data, was 'pulled out' for consideration here.

Method

Subjects. Eight paid undergraduate volunteers (four men, four women) were employed. They met the previous criteria for good vision.

Materials. The eight matrices of Experiment I, Categories 1 and 2, were used again. These represented the eight possible combinations of size, element-quadrant relationship, and figure-ground relationship.

Apparatus. The apparatus was the same as before.

Procedure. The procedure was kept as similar to that of Experiment III as possible. Each *S* was run in a single session of about 1½ hr., including 128 trials (8 matrices × 8 rotations × 2 head positions). These were divided into eight blocks of 16 trials each, with head position alternated every block and size alternated every two blocks. Initial size and head position were counterbalanced over *S*s.

Results

A four-way analysis of variance yielded two significant main effects and one significant second-order interaction, as noted below; no other effects approached significance. The size of the matrix showed no importance either as a main effect or in interaction with other variables. Element-quadrant relationship did give a moderately significant main effect [F (1, 7) = 7.0, $p < .05$], but no significant interactions. The *S*s located the disparate quadrant more quickly when its boundaries were in a parallel-perpendicular relationship to elements (1.06 sec.) than when the boundaries and elements were in a 45° relationship (1.30 sec.).

The major results, involving element orientation and head orientation, are summarized in Table IV. The main effect of element orientation was highly significant [F (1, 7) = 33.8, $p < .001$]. This is the effect attributable to *physical*, or *gravitational*, orientation of elements, with retinal orientation balanced out: the mean RTs compared were 1.09 sec. for stimuli with physically horizontal and vertical elements, and 1.28 sec. for stimuli with physically diagonal elements (see Table IV). The interaction of

TABLE IV

EFFECTS OF HEAD AND LINE-ELEMENT ORIENTATION: EXPERIMENT IV

	Head orientation			
	Head upright	Head tilted	Mean	Diff.
Element orientation				
HV	.89	1.29	1.09	−.40
D	1.51	1.04	1.28	.47
Mean	1.20	1.17		
Diff.	−.62	.25		

element orientation with head orientation was also significant [$F(1, 7) = 25.0, p < .01$]. This is the effect attributable to *retinal* orientation, with gravitational orientation balanced out: the means involved were .97 sec. for stimuli with retinally horizontal and vertical elements, and 1.40 for stimuli with retinally diagonal elements. (These values were obtained by averaging the diagonal cells of Table IV. Note that the designations 'main effect' and 'interaction' are quite arbitrary and would be reversed if the elements were described in terms of retinal orientation instead of physical orientation.) A paradoxical aspect of these results is that the effect of physical orientation was more significant than that of retinal orientation, although the effect of retinal orientation on mean RT was more than twice that of physical orientation. This means that the physical orientation effect, though smaller, was much more consistent. It is obvious that the results of this experiment were in good agreement with those of the experiment with paired dots.

DISCUSSION

Slope was by far the most powerful of the grouping variables studied in Experiment I, and we saw further that some effects superficially associated with higher-order variables could probably be attributed to slope in one sense or another. In Experiments II, III, and IV, the question was whether slope, as it contributes to grouping, should be defined in terms of retinal or physical coordinates. The more basic question is whether similarity grouping occurs before or after the achievement of orientation constancy, or possibly at both levels. This question still cannot be answered completely or unequivocally. At the very least, the significant physical orientation effects found in Experiments III and IV seem to rule out the possibility that grouping occurs entirely at the projection level, *e.g.* among slope analyzers rigidly connected to the retina. These effects are consistent with the hypothesis (discussed above) that grouping occurs among descriptors based on relationships to a labile Cartesian reference system. On the other hand, retinal orientation *also* had very important effects in both Experiments III and IV; and in Experiment II, which employed angles instead of line segments, *only* retinal orientation was important.

Two alternative approaches to the explanation of these results may be suggested. The first, and the most obvious possibility, is that horizontals and verticals receive favored treatment *both* at a projection level *and* at a higher level where phenomenal orientation is represented. (This is not to say that grouping processes occur at both levels; they could be confined to the higher.) The assumption here is that phenomenal orientation is identifiable with physical orientation when the head is tilted. This is not necessarily true: Attneave and Reid (see n. 4) found that their subjects *could* adopt a head-oriented reference system when instructed to do so. However, the subjects in the present experiments were explicitly told to think of the physical top of the display as 'up.' A second, alternative possibility is that the horizontal and vertical have special importance *only* at the higher level but that the phenomenal references axes are during some intervals head-anchored, during others gravity-anchored. The momentary probability of one reference orientation as opposed to the other might depend upon various task and stimulus parameters, as well as upon instructions. One such factor might be the complexity of the stimulus situation: note the reversal in importance of physical and retinal orientation between the discriminative reaction-time studies of Attneave and Olson (see n. 4) using single line segments and the present Experiment II employing the most complex material studied. In any event, a tenable explanation must take some account of constancy mechanisms.

LEARNING AND RETENTION OF LETTER PAIRS AS A FUNCTION OF ASSOCIATION STRENGTH

HARRIETT AMSTER, GEOFFREY KEPPEL, and ANN MEYER,
University of California, Berkeley

Abstract. Three experiments on paired-associate learning and retention of letter pairs tested the Underwood-Postman hypothesis that extraexperimental associations interfere with learning and produce forgetting. As hypothesized, in experiments with seven- and ten-year-old children and with adults, highly associated (high-strength) letter pairs were learned faster and, on the whole, retained better over one week than low-strength pairs. The positive retention effects, not observed in previous studies, are partly attributed to the use of the correction procedure as a control for the overlearning of easy items in low-strength lists that is unavoidable with standard methods. This overlearning is believed to reduce the probability of detecting any differential retention of high- and low-strength lists. Developmental differences in children were small for original learning and insignificant for forgetting. A single postcriterion trial, however, facilitated retention-test performance. Theoretical implications are discussed.

The most explicit version of interference theory accounts for the forgetting of verbal materials in terms of extraexperimental or linguistic interference.[1] This interference could operate between words (unit-sequence interference) or between letters (letter-sequence interference). In both cases, it is assumed that the major source of interference responsible for the forgetting of a particular association stems from responses which are more strongly associated extraexperimentally with the stimulus term than is the prescribed response. One expectation of the theory, then, is an inverse relationship between association strength and forgetting, since weak responses have more potential competitors than strong responses.

Received for publication September 9, 1969. The research was supported by grants (HD-1145 and MH-10249) from the NIH and conducted at the Institute of Human Learning, which is also supported by a grant from NIH. The senior author conducted the research under a Research Career Development Award from NIHD.

[1] J. Underwood and L. Postman, Extraexperimental sources of interference in forgetting, *Psychol. Rev.*, 67, 1960, 73–95.

There have been several tests of this prediction for both sources of interference. Tests of unit-sequence interference have generally supported the theory. For example, Postman found a greater decrement in the recall of weakly associated (low-strength) word pairs, but this was true only for stimulus terms of low frequency.[2] Higgins replicated these findings, although it is not certain that the different sets of pairs were learned to an equal degree.[3] In a later study, Postman, Fraser, and Burns obtained better recall of highly associated than of unrelated word pairs.[4] Finally, Turnage, using a multiple-choice retention test, reported greater forgetting of low-strength word pairs, an effect that was subsequently replicated with letter-word pairs by Turnage and Anderson, using the same testing procedure.[5]

Tests of letter-sequence interference, however, have *not* generally supported the theory. Underwood and Keppel, using the letter-association norms of Underwood and Schulz, selected single-letter pairs representing two levels of association strength.[6] Although the two sets of pairs were learned at different rates, there was no evidence of differential forgetting at any of the points at which unambiguous comparisons could be made. In a subsequent experiment, Underwood again failed to find a difference in forgetting letter pairs, even when these were more widely separated in normative association strength and showed an approximately fourfold difference in learnability.[7] Montague and Wearing also reported equal forgetting of lists of high- and low-strength letter pairs, but their results are difficult to interpret because of differences in original learning of the lists.[8]

---

[2] L. Postman, The effects of language habits on the acquisition and retention of verbal associations, *J. exp. Psychol.*, 64, 1962, 7–19.

[3] J. Higgins, Effect of response probability on acquisition and retention of paired associates with varying stimulus frequency: Confirmatory results, *J. verb. Learn. verb. Behav.*, 6, 1967, 446–447.

[4] L. Postman, J. Fraser, and S. Burns, Unit-sequence facilitation in recall, *J. verb. Learn. verb. Behav.*, 7, 1968, 217–224.

[5] T. Turnage, Pre-experimental associative probability as a determinant of retention, *J. verb. Learn. verb. Behav.*, 2, 1963, 352–360; T. Turnage and N. S. Anderson, Letter-frequency and associative probability as determinants of learning and retention, *J. verb. Learn. verb. Behav.*, 4, 1965, 463–468.

[6] B. J. Underwood and G. Keppel, Retention as a function of degree of learning and letter-sequence interference, *Psychol. Monogr.*, 77, 1963 (No. 4), 1–16; B. J. Underwood and R. W. Schulz, *Meaningfulness and Verbal Learning,* 1960, 1–430.

[7] B. J. Underwood, Degree of learning and the measurement of forgetting, *J. verb. Learn. verb. Behav.*, 3, 1964, 112–129.

[8] W. E. Montague and A. J. Wearing, Natural language mediation: A source

EXPERIMENTS I AND II

The present studies were conceived as a further test of the hypothesized letter-sequence interference and as a vehicle for studying developmental changes in associative mechanisms. Many of the reasons for the negative results on differential retention of high- and low-strength letter pairs could be less applicable to children than to adults. For example, a major difference between adults and children in associative responding to letters is the increase with age in the tendency to make meaningful units, *e.g.* word-completion responses or institutional abbreviations, from letter stimuli.[9] To the extent that adults may employ such a strategy during the course of learning letter pairs, and to the extent that such strategies may be effectively applied to both high- and low-strength pairs, any differential susceptibility to forgetting such pairs could be reduced in adults to a greater degree than in children. In addition, when a child begins to read and spell, he may be especially susceptible to letter-sequence interference and, as a consequence, show greater forgetting of low-probability letter pairs. But as he masters these skills, he may 'process' words as units rather than as a sequences of letters, thus diminishing the relative importance of letter-sequence interference. Thus, predictions that rate of forgetting low- and high-strength letter pairs interacts with age could, if verified, account for previous failures to observe differential forgetting of letter pairs with adults.

Furthermore, although association hierarchies for letters show a developmental increase in steepness when estimated from the increase in the strength of the primary and other dominant responses, there is a parallel developmental increase in the number and frequency of very rare responses (see n. 9). Although it is not known whether the number or strength of competing responses is the critical source of forgetting, it is clear that the available number of low-strength competitors of a low-strength response should be greater for adults than for children. On this basis, forgetting of weak responses could increase with age as the number of weak competitors increases, implying better retention of low-strength letter pairs for children than for adults. But on the other hand,

---

of interference with extra-experimental interference, *Psychon. Sci.*, 9, 1967, 317–318.

[9] H. Amster and G. Keppel, Letter-sequence habits in children, *J. verb. Learn. verb. Behav.*, 7, 1968, 1326–332.

any differences in letter-pair retention between, say, children of seven years of age and children of ten may be related to the increasing strength with age of dominant responses rather than weak competitors, since the increase in number and frequency of rare responses between those ages is negligible. Because the use of letter pairs permits the assumption that the number of possible responses is equivalent for all literate *S*s, any age differences in learning and retention attributable to associative factors may be said to result from changes in the shapes of the hierarchies (and not from concomitant changes in vocabulary size, as occurs with words). Also, the constant size of the pool of letter responses simplifies developmental comparisons by unconfounding frequency as a measure of association strength, with age changes in the size of the pool.

A major requirement of any retention study is to bring different treatment groups to the same level of learning before the retention interval in order to separate the effects of the variables under investigation from those produced by differences in original learning. Underwood has discussed and evaluated two means by which such an equation may be accomplished: (*a*) by matching groups through the administration of differential numbers of trials, the exact numbers of which are based upon a consideration of learning curves obtained from a pilot study; and (*b*) by matching *S*s through the use of a common performance criterion (see n. 7). The first procedure was used in an unpublished study on the forgetting of weakly and highly associated letter pairs by second- and fifth-grade children.[10] Lists ('low' and 'high') were drawn both from adult norms and from norms relevant to the ages of the *S*s.[11] On the basis of learning curves obtained from 10 *S*s per group (eight groups), estimates were made of the numbers of trials needed to produce a learning criterion of 5/7 correct. For instance, on the age-relevant lists second graders attained this criterion following Trial 13 on the low list and Trial 7 on the high list, while on the lists based upon the adult norms the values were 17 and 13 trials respectively. The main experiment was then conducted with the appropriate numbers of adjusted trials.

The results revealed two significant difficulties. First, the equation of groups was inadequate for our purposes. Out of the four comparisons of learning on low and high lists, only one was sufficiently close to allow a reasonable comparison of retention scores. The other three deviated from .90 to 2.70 correct at the end of learning. Other investigators have been far more successful with this sort of equating procedure,[12] but all of their studies have been conducted with college students. Perhaps the difference in success derives from the possibility that samples of college students are more homogeneous in learning ability than are samples of children drawn from the public schools. It is conceivable that larger samples and a better matching of subject pools for the pilot and main experiments would solve this difficulty in future attempts.

---

[10] The authors wish to thank A. Gordon, who collected and analyzed the data for this unpublished study.

[11] For the Underwood-Schulz adult norms, see n. 6 above; for the age-relevant norms, see H. Amster and G. Keppel, Letter association norms, *Psychon. Monogr. Suppl.*, 1, 1966 (No. 9), 211–238.

The second difficulty with this approach was reflected in the amazingly high level of retention observed one week after learning. Forgetting ranged from 2.4% to 37.5%, with an average of 23.9%. While direct comparisons are not possible due to various differences in procedure, adults in the Underwood-Keppel study (see n. 6) averaged approximately 51% forgetting over the same time period if equations are made for estimated terminal levels of learning. One likely explanation of these different amounts of forgetting is the strong possibility of hidden overlearning for the children. Fatigue, waning attention, and boredom can affect their performance and thereby mask overlearning. Whereas such discrepancies between learning and performance are not usually suspected in verbal-learning research with adults, they are considerably more probable with children. Thus, the trial averages, which formed the basis for the adjusted numbers of trials in the main retention study, may have seriously underestimated the actual level of learning attained and resulted in an artificially high level of recall when the deleterious effects of practice were no longer present a week later.

As a result of these considerations, changes were made in the equating procedure and in the method of learning for the experiments that follow. *S*s were taken to a common performance criterion, and instead of the traditional anticipation method of learning, a correction method of learning was used.[13] With this method *S*s are forced to respond correctly to each stimulus in the list before being given further practice on the list as a whole. The value of this procedure was seen as twofold: it may reduce differential overlearning between conditions and among items within conditions, and it might maintain a high level of motivation for the young *S*s. In Experiment I the learning and retention of the low and high lists were contrasted for second- and fifth-graders; additional second-graders were given an immediate postcriterion test in order to evaluate the success with which the groups were equated in terms of terminal degrees of learning. Lists were selected from the relevant children's norms because the unpublished work had revealed a more predictable relationship between association strength and learning under these circumstances than when the lists were selected from adult norms. Experiment II replicated Experiment I, but with the immediate-retention groups included for both grades.

---

[12] *E.g.* Postman, Fraser, and Burns (see n. 4 above) and B. J. Underwood and B. R. Ekstrand, Linguistic associations and retention, *J. verb. Learn. verb. Behav.*, 7, 1968, 126–171.

[13] See W. F. Battig, Procedural problems in paired-associate learning, *Psychon. Monogr. Suppl.*, 1, 1965 (No. 1), 1–12.

Method

Design. The design of Experiment I was an incomplete 2 × 2 × 2 factorial, with association strength (low and high) and age (7 and 10 yr.) as the two completely crossed variables, and with retention interval (immediate and delayed for second graders; delayed only, for fifth graders) as the incomplete variable. Experiment II was a complete factorial involving the same variables. In all other respects Experiments I and II were identical.

Materials. Two lists, with seven pairs of letters in each, were selected from age-relevant letter-association norms (see n. 11) for the two age-groups. The second-graders' low list was F-K, C-J, Y-G, M-D, V-L, A-Q, K-P; the second-graders' high list was F-I, C-A, Y-R, M-O, V-S, A-B, K-E; the fifth-graders' low list was I-V, K-F, X-D, U-X, W-C, N-B, Q-G; the fifth-graders' high list was I-O, K-E, X-Y, U-D, W-A, N-M, Q-U. The high lists had average percentage forward association strengths of 12.9% and 13.6% for second- and fifth-graders respectively; the low lists, averages of .4% and .5%. (These values refer to the percentages of *S*s giving these particular letter associations on the letter-association norms.) Within each age-group, the same stimulus terms were used, so that the low and high lists were differentiated entirely in terms of the response term. In each list one letter was used both as a stimulus and a response term. In selecting letter pairs, an attempt was made to hold backward association strengths constant at a low value. The respective percentages were 3.9% and 3.8% for the two high lists, and 2.0% and 2.5% for the two low lists.

Procedure. The *S*s were trained to criterion using the correction method basically as described by Battig (see n. 13). The training began with a *main trial* consisting of a study trial and then a test trial on all seven pairs, followed by a *subtrial* consisting of a study trial and a test trial on those items that had not been given correctly on the preceding main trial. Subtrials were repeated in the same manner by continually dropping out pairs attained on the preceding subtrial, until each pair was correctly anticipated once. At this point a new main trial and new subtrials were given. This procedure was repeated until *S* achieved a criterion of 6/7 correct on a *main* test trial.

Five different orders of the pairs were used on the study trials. The test-trial stimuli were also arranged in five orders which differed from the study orders. Starting position was selected at random for each *S*, but each starting position was represented equally within each group. The method of presentation entailed the use of the designated starting order on the first main trial and successive predesignated orders on the subsequent main trials or subtrials. On most subtrials *E* could merely delete from the prearranged order those items that had been attained. When only a few items remained, it was occasionally necessary to reverse the order to prevent more than two successive presentations of two stimuli in the same order.

The materials were presented by means of a Kodak Carousal projector controlled by a Time-O-Lite timer. The study and test intervals were both set at 2 sec., with an interstimulus interval of approximately 3 sec. The intervals between each study-test sequence, whether a main trial or a subtrial, was manually controlled and minimized to about 3 sec.

Instructions on the correction method were supplemented by a practice task consisting of five pairs of pictures of common objects, *e.g.* cat, house, boat, bell, etc. The practice task was taken to a criterion of 4/5 correct before the experimental task was undertaken. Following the attainment of criterion on the main task, *S*s in the immediate groups were given an additional test trial, while training was terminated for *S*s in the delay groups. All *S*s were given a retention test after 1 wk. (A comparison of immediate and delay groups on the delayed test allowed a determination of the influence of the immediate-retention test on delayed recall.) All of the stimuli were presented on each retention test. The testing procedure was identical to that used in learning.

Subjects. A total of 60 children served in Experiment I, with 10 *S*s in each basic list-grade-interval combination. The *S*s were drawn from the University of California Demonstration Summer School at Whittier School in the Berkeley Unified School District (tuition required). There were 80 *S*s in Experiment II, 10 in each of the eight treatment conditions. These *S*s were attending Whittier School and were tested during the normal tuition-free school year. Both samples of *S*s were roughly of middle-class socioeconomic status, although the summer-school sample was probably higher in status but more variable in ability. The *S*s were assigned randomly to conditions in blocks of four (Grade 2) and two (Grade 5) in Experiment I and in blocks of four in both grades in Experiment II. Only one *S* was dropped (from Experiment I, because of absence for the retention test), and he was replaced.

Results

Learning. With the correction method of learning, typical measures of acquisition such as trials to criterion do not provide a sensitive index of learning differences. This is a consequence of the fact that the 'remedial' training given slow *S*s or difficult pairs tends

TABLE I

MEAN TOTAL EXPOSURES PER PAIR

	Association strength					
	Low			High		
	Immediate	Delay	Mean	Immediate	Delay	Mean
Experiment I†						
Grade 2	6.87	8.81	7.84	4.54	4.98	4.76
Grade 5		6.88	6.88		3.36	3.36
Experiment II‡						
Grade 2	5.84	8.94	7.39	5.01	4.37	4.69
Grade 5	6.90	5.18	6.04	2.64	3.33	2.98

† MS—within = 9.484.
‡ MS—within = 7.335.

to reduce differences among groups in rate of acquisition. The measure used here counted the total number of pair exposures needed, on main and subtrials, to reach criterion. These scores, divided by the number of pairs, are presented in Table I. In both experiments there was a sizable effect of association strength ($p < .001$), low-strength pairs requiring more presentations than high-strength pairs. Age effects were small, fifth-graders tending to learn more rapidly than second-graders. This difference was not reliable in Experiment I [F (1, 54) = 1.96, $p > .10$] but attained an acceptable level of significance in Experiment II [F (1, 72) = 6.37, $p < .025$]. In neither experiment was there an indication of any interaction of age and association strength (Fs < 1). (It should be noted that any comparison between grades must be interpreted with caution since different lists were used at each grade level. On the other hand, the two sets were matched on all known relevant variables so that this confounding of grades and materials was not considered serious.)

Where available, immediate and delay groups were compared for speed of learning. In all cases except one, these differences were not significant (ps > .10). The exception occurred in Experiment II, for the second-graders learning the low list [F (1, 72) = 6.55, $p < .025$]. There is no explanation for this difference other than the vagaries of random sampling. This inequality is probably not serious for the retention part of the study since all *S*s were carried to the same performance criterion and the correlation between total presentations and recall was essentially zero.

Recall. Table II presents a number of measures useful to interpretation of the retention data. The first column lists the average performance of the different groups on the criterial trial. An inspection of these values reveals relatively homogeneous performance of the 14 groups, indicating that in spite of differences in learning, the groups learning the low and high lists ('low' and 'high' groups) tended to approach the same terminal level of learning. As Underwood has argued, however (see n. 7), this does not guarantee that *S*s actually reached the same level of learning. The immediate-retention data provided a check on this. A comparison of immediate and criterial scores shows a loss but still fairly homogeneous levels for the low and high groups ($\bar{X} = 5.97$ and 6.10 respectively). In terms of percentages, these scores represent a 5.7% drop for the low list and a 3.6% drop for the high list. In short, these data indicate that the groups were well matched in terminal performance and that any differences present may be 'taken care of' by considering loss scores.

The recall scores obtained following the one-week retention interval are given in the third column of Table II. The overall level of retention was quite comparable to that obtained with adults who learned by the anticipation method. With immediate recall as a

TABLE II

MEAN RECALL AND LOSS SCORES

	Criterial scores		Immediate recall		Delayed recall		Loss scores	
	Low	High	Low	High	Low	High	Low	High
Experiment I								
Grade 2								
Immediate	6.4	6.3	5.7	5.9	3.1	3.8	2.6	2.1
Delay	6.0	6.4			2.8	3.1	3.2	3.3
Grade 5								
Delay	6.4	6.2			3.4	3.8	3.0	2.4
Experiment II								
Grade 2								
Immediate	6.3	6.4	6.4	6.0	3.7	4.6	2.7	1.4
Delay	6.2	6.4			2.6	4.2	3.6	2.2
Grade 5								
Immediate	6.3	6.3	5.8	6.4	3.6	4.5	2.2	1.9
Delay	6.2	6.2			3.1	3.7	3.1	2.5

baseline, the average level of retention for the combined delay groups was 55%. It was noted earlier that college-student *S*s recalled 51% of the letter pairs over the same retention interval (see n. 6). Thus, it appears that the considerably higher percentage obtained previously (in the unpublished study mentioned above) was not the result of a greater resistance to forgetting for children but, rather, attributable to overlearning of the pairs. In the main comparisons of interest, there was, here, consistently higher recall of the high-strength lists, a difference predicted by the Underwood-Postman theory (see n. 1). In terms of percentages (delay/immediate), retention was 49.9% for the low list and 60.6% for the high list. The same pattern is present when one considers the delayed recall of the immediate groups.

Statistical analyses were complicated, since immediate-retention groups were not included for the fifth graders in Experiment I. This problem was solved in two ways. First, immediate and delayed retention intervals were employed in two analyses: Experiment I, Grade 2 only; and Experiment II, both grades. For Experiment I the analysis was a 2 × 2 factorial (strength and interval). The main effect of strength and the critical interaction were not significant [the Fs $(1, 36) = <1$], but there was reliable forgetting [$F\ (1, 36) = 67.13$, $p < .01$]. For Experiment II the analysis was a 2 × 2 × 2 (strength, interval, and grade). This analysis also showed reliable forgetting [$F\ (1, 72) = 144.46$, $p < .01$] and, moreover, superior recall of the high lists [$F\ (1, 72) = 6.68$, $p < .025$]. The critical interaction of interval and strength was significant, but since the interaction of interval, strength, and grade was also significant [$F\ (1, 72) = 4.78$, $p < .05$], the nested effect was assessed. For Grade 2, the critical interaction was highly significant [$F\ (1, 36) = 9.55$]; but not for Grade 5 ($F < 1$). Thus, although the means were in the direction of superior recall for the high lists in all groups for both experiments, it was only for Grade 2 in Experiment II that the effect was both significant and supported by a significant critical interaction.

A second analysis compared every available low-high comparison independently. For this analysis, loss scores were used to correct for any differences in degree of learning. For the immediate groups, this latter estimate came from the immediate test; for the delay groups, it came from the criterial test. The actual values entering into this analysis are shown in the last column of Table II. The average loss scores were 2.91 and 2.26 for the low and high lists respectively. An analysis of variance, treating the seven comparisons as one factor and associative strength as the other, produced a significant difference for the low-high comparison [F (1, 126) = 7.63, $p < .01$]. None of the other sources was reliable (ps $< .10$). The same conclusion is reached if the seven comparisons are treated as independent tests of the same hypothesis [X^2 (14) = 29.88, $p < .01$].[14]

A final comparison of interest was between the immediate and delay groups on the delayed-retention test. In all six of the possible comparisons, recall was higher for the immediate groups. An analysis of Experiment II, where the factorial design was completely balanced, substantiated this observation [F (1, 72) = 5.36, $p < .025$]. None of the interactions involving this variable was significant (Fs < 1.06). From these data, then, it is clear that a single postcriterion recall trial can considerably enhance recall on a delayed test. This conclusion runs counter to that suggested by other studies concerned with this effect: Montague and Kiess, for instance, employed the correction method with adult *S*s and yet failed to find an effect on retention of three repeated tests.[15] The criterial differences between such studies and the present ones are difficult to isolate a priori; nevertheless, they are important to future research with children on the effect of postcriterion test trials upon delayed recall, for, as has been seen, the effect can be quite large.

EXPERIMENT III

Why did degree of letter-sequence interference influence forgetting in children when it has not done so for adults? One strong possibility is that the present experiments with children used the correction procedure, while the experiments with adults used standard anticipation learning (see nn. 7 and 9). An advantage of the correction method is that it may reduce differential overlearning among pairs, which may account for the failure to find results with the anticipation method. Although overlearning of individual pairs is a problem in any retention study, greater differential learning can occur among pairs within difficult lists than among pairs in easy lists. The differential overlearning comes about in any list by virtue of the fact that the first-learned pairs receive overlearning trials while the remaining pairs are acquired. Moreover, the extent to which this occurs tends to increase with the difficulty of the list. Thus, greater forgetting of low-strength pairs in the adult studies may have been offset by greater overlearning of easy pairs in the low list. To the extent that the correction method can reduce differential overlearning, it should provide a method for revealing retention differences between high and low lists and thus support the Underwood-Postman theory with college-age *S*s.

---

[14] B. J. Winer, *Statistical Principles in Experimental Design,* 1962, 43–45.

[15] C. N. Cofer, F. Diamond, R. A. Olsen, J. S. Stein, and H. Walker, Comparison of anticipation and recall methods in paired-associate learning, *J. exp. Psychol.,* 75, 1967, 545–558; W. E. Montague and H. O. Kiess, Test of a procedure to control inter-pair learning and the effect of repeated, post-criterion test trials on retention, *Psychon. Sci.,* 6, 1966, 191–192.

Method

Design and materials. There were two basic groups, and the major experimental variable was the association strength of the letter pairs. To reduce the speed of learning the high list, the number of letter pairs was increased from 7 to 11. Four different lists at each level of strength were selected from the Underwood-Schulz adult norms (see n. 6). The lists were constructed in four low-high sets (see Table III); for each set the same 11 letters constituted the stimuli while different responses determined the association strength. No letter was used more than once in any list. The average association strength for the low list was 1.5%; for the high list it was 15.2%.

Procedure. Essentially the same correction procedure used in Experiments I and II was followed in Experiment III. Verbal instructions were abbreviated

TABLE III

LISTS FOR EXPERIMENT III

	Set 1		Set 2		Set 3		Set 4	
	Low	High	Low	High	Low	High	Low	High
	B–L	B–E	Z–S	Z–T	Q–F	Q–U	X–S	X–Y
	C–G	C–D	J–T	J–K	R–D	R–S	K–Y	K–L
	X–F	X–Y	H–K	H–I	V–P	V–W	M–V	M–N
	J–N	J–P	L–V	L–M	H–S	H–E	W–D	W–U
	Q–Z	Q–T	R–M	R–S	X–W	X–I	Z–P	Z–T
	R–U	R–S	W–P	W–E	C–N	C–O	G–N	G–O
	W–P	W–U	Q–N	Q–U	Y–B	Y–A	Q–F	Q–R
	H–S	H–O	B–F	B–O	J–T	J–P	H–U	H–I
	I–A	I–N	X–U	X–Y	Z–L	Z–T	B–R	B–E
	K–Y	K–L	C–E	C–D	K–G	K–L	C–L	C–A
	V–M	V–A	G–D	G–A	M–U	M–N	J–T	J–P
Mean strength	1.8%	11.5%	1.6%	17.0%	1.1%	13.6%	1.5%	18.6%

and the presentation rate was increased, however. For Experiment III the pairs were exposed for 1 sec. and the testing rate was set at 1.5 sec. per pair. Intertrial intervals were controlled manually and minimized to about 3 sec. The *S*s were taken to a criterion of 9/11 correct. Following the criterial trial all *S*s were dismissed; they were asked to return for a recall test 1 wk. later but were not informed of the purpose for their return.

Subjects. The *S*s were 40 undergraduates in psychology courses at the University of California, Berkeley. They were assigned randomly to low-high lists in blocks of 4 *S*s; lists were balanced over the course of the experiment, 20 *S*s per basic condition and 5 *S*s per list within each condition. No *S*s were excluded for failure to return for the delayed test.

Results

Learning. In spite of a generally rapid rate of acquisition, a difference still emerged between the two conditions. The mean numbers of presentations per pair to criterion were 3.43 and 2.23 for the low and high lists respectively [F (1, 32) = 24.12]. In addition, list sets produced significant variability [F (3, 32) = 3.53, $p <$ .05]. The fact that the interaction of sets and strength was not reliable [F (3, 32) = 2.00, $p >$ .10] indicates that strength had a general effect and that the differences among the sets may have been due to the particular stimuli chosen for each set. Inspection of the lists did not reveal any obvious stimulus differences, however.

Recall. As in Experiments I and II, the low list was more poorly recalled than the high. The averages were 4.75 and 7.30 respectively. This difference was unaffected when loss scores were computed, since identical criterial scores (10.25) were obtained in the two conditions. An analysis of the loss scores showed the differential forgetting of the two lists to be significant [F (1, 32) = 17.17]. Again, differences among the list sets were reliable [F (3, 32) = 10.46, $p <$ .01], while the interaction was not ($F <$ 1). There was no strong relationship between list difficulty and amount lost, an ordering from the easiest to the most difficult set being associated with a 2-1-4-3 ordering from the smallest to the greatest loss. As a further exploration of the effect of learning speed on forgetting, *S*s were divided on the basis of the total number of presentations to criterion. There was some evidence of a smaller loss for the fast *S*s (mean loss = 3.59 and 5.00 for the fast and slow learners respectively), but neither this difference, nor the interaction of strength and subject, was reliable [F (1, 36) = 3.16 and < 1 respectively; ps > .05]. In short, we were unable to discover the factor(s) contributing to the differences in learning and recall for the four stimulus sets; but the fact that the sets did not interact with the strength variable allows the general conclusion that high-strength letter pairs were learned faster and forgotten more slowly than were low-strength pairs.

DISCUSSION

The learning data of Experiments I and II confirm other reports that association norms can provide fairly accurate prediction of the rates at which children learn particular letter pairs in the laboratory.[16] The data of Experiments I and II (and III) further show that even relatively small differences in assumed association strength produce sizable differences in learning. The only unambiguous developmental comparisons showed only small overall age differences in learning between second- and fifth-graders—not a surprising result, considering that only three years separate the two grades and that age differences have not been found with separations of two to four years.[17] More important was the possibility of finding an interaction of association strength and age, for earlier research on *word*-association pairs suggests that the influence of association strength decreases for the older *S*s. In neither Experiment I or II, however, was there any indication of such an interaction of association strength and age. Moreover, it is highly unlikely that even considerably older subjects would show an elimination of the difference in the learning of low- and high-strength letter pairs. For instance, the college students in Experiment III still exhibited a reliable difference. In addition, Underwood reported substantial differences with college students learning lists of approximately the same levels of strength (0% and 13.9%; see n. 7) as those employed in Experiments I and II (0% and 12.9%–13.6%). Thus, it is possible that the influence of association strength

[16] *E.g.* J. C. McCullers, An analysis of some factors underlying intra-list associate transfer in paired-associate learning *J. exp. Psychol.*, 65, 1963, 163–168; also, D. A. Wicklund, D. S. Palermo, and J. J. Jenkins, The effects of associative strength and response hierarchy on paired-associate learning, *J. verb. Learn. verb. Behav.*, 3, 1964, 413–420.

[17] J. C. McCullers, Effects of associative strength, grade level, and interpair interval in verbal paired-associate learning, *Child Develpm.*, 32, 1961, 773–778; S. S. Shapiro, Paired-associate learning in children: Length of list, initial associative strength, presentation time, and grade level, *Psychol. Rep.*, 20, 1967, 903–908.

exhibits a different developmental trend in the learning of word pairs than it does in the learning of letter pairs.

An important source of difference between investigations using letter pairs and those using word pairs is that the latter often omit an important control: common stimulus terms in high and low lists. The use of common stimuli ensures that for every pair in the high list there is a pair from the same hierarchy in the low list which is definitely lower in associative strength. (This is so even though the association strengths may not conform precisely to their frequencies.)[18] Since no evidence for developmental differences in the effects of association strength was here observed, it is not possible to speculate about the relative importance of the size and shape of the hierarchy in accounting for retention. However, the overall superiority of learning and retention with the high list can certainly be attributed to association strength per se rather than to either size or shape of hierarchy, since the high and low pairs came from the same hierarchies in every case. Although this fact is believed to contribute greatly to the possibility of revealing differences in retention due to associative strength, it is noted that Underwood's letter pairs were controlled in this manner also; yet he failed to obtain a retention effect (see n. 7).

Perhaps the age difference in trends for *word* pairs can be explained by a consideration of the cumulating frequency with which particular pairs are experienced in the course of normal linguistic activity. Assuming that an asymptote is reached in the tendency of additional exposure to increase 'learnability,' low-strength pairs will eventually 'catch up' with high-strength pairs. At this point, word-association strength will no longer influence speed of paired-associate learning. Essentially the same explanation has been offered by Wicklund *et al.* (see n. 16). Using this 'asymptote' explanation, certain other facts fall into place. For example, it is less likely that word associates of infrequent English words reach asymptote; as a consequence, differences in learning as a function of association strength can still be obtained with college students, as Postman reported (see n. 2). Similarly, the fact that word pairs drawn from restricted association norms predict the learning of college students[19] can be explained by assuming that high-strength pairs from these norms have not reached asymptote either. This assumption is not unreasonable in the light of the difficulty of obtaining such norms.[20]

---

[18] The shape of the hierarchy as well as the frequency appears to affect learnability, as shown by Wicklund, Palermo, and Jenkins (see n. 16 above).

[19] B. J. Underwood and R. W. Schulz, Response dominance and rate of learning paired associates, *J. gen. Psychol.*, 63, 1960, 153–158.

In contrast, *letter* pairs are probably not affected in the same way by increased experience with the language. That is, high-strength pairs for adults represent the production of alphabetical sequences and the continuation or completion of well-integrated letter habits, *e.g.* words and common abbreviations.[21] While high-strength letter pairs undoubtedly reach an asymptote of strength early, low-strength pairs may show a slow and gradual increase. Unlike low-strength word pairs, low-strength letter pairs receive little systematic linguistic input and thus remain far lower in absolute preexperimental strength than pairs which appear in words. This interpretation of the present results does not account for the failures of Underwood and Keppel and of Underwood to find retention differences for adults (see nn. 6 and 7). It is, rather, based on notions that require experimental verification, *i.e.* developmental investigations of the influence of association strength on learning for a variety of association norms.

The retention data of Experiments I, II, and III provide considerable support for the Underwood-Postman theory of extra-experimental interference. Admittedly, the differences in the recall of low and high lists were not great for the children in Experiments I and II, but they were consistent. It is conceivable that the differences could be magnified with greater contrast in association strength, or that there is a genuine increase in the effect with age. Age differences in forgetting were not observed, but perhaps they should not be expected since (*a*) the explanation above would predict an early asymptote in effective strength, (*b*) age produced only small effects in learning, and (*c*) speed of learning was not consistently related to retention.

The college-age *S*s in Experiment III were tested to determine whether differential forgetting of low and high pairs might be observed when the correction method was used during acquisition, thus accounting for earlier failures to find retention differences as a function of associative strength. The answer to this question was impressively affirmative, retention being 46.3% for the low pairs and 71.2% for the high pairs. Since immediate-retention groups were not included in this experiment, there is the possibility that the low and high lists were not at the same level of learning at the end of the criterial trial. While it is true that the groups were equal on this trial, one might still expect the high group to benefit more from the terminal recall trial than would the low group. Two arguments can be offered against this possibility, however. First, the immediate-retention groups of Experiments I and II showed that an extra recall test (the postcriterion trial) can enhance retention but that this gain was equivalent for the low and high lists. There is no reason to believe that this would not also have been true in Experiment III. Second, even if a small difference in degree of learning were present, it is unlikely that this discrepancy could account for a difference as large as that observed on the delayed test. We conclude, therefore, that the differential forgetting of low and high pairs was due to factors coming into play following learning and not to a difference in degree of learning.

---

[20] *E.g.* B. J. Underwood and J. Richardson, Some verbal materials for the study of concept formation, *Psychol. Bull.*, 53, 1956, 84–95.

[21] See Amster and Keppel (n. 9 above) and Underwood and Schulz (n. 6 above).

One reason for the success of this experiment and the failure of others has already been mentioned, namely, an overlearning of the easier pairs in the low list. There is no corroborating evidence in support of this speculation, nor have direct comparisons of the correction and noncorrection procedures in learning and retention as yet been reported. Until such comparisons are made, we can only suggest that the correction method may offer a more sensitive method for assessing differences in forgetting and that this increased sensitivity involves a reduction of interpair and intersubject variability.

As detailed in the introduction to this paper, the greater forgetting of the low-strength letter pairs can be interpreted in terms of an increased amount of letter-sequence interference relative to that present for high-strength letter pairs. It might also be interpreted in terms of response generation. That is, if *S*s give high-strength responses at recall when they cannot produce the correct responses, they may be given credit for 'hits' when in fact they were merely guessing. This, of course, is a difficult proposition to reject, although we can offer three comments which tend to minimize its relevance to the present experiments. First, guessing was not widespread since there were few overt errors at recall, and they occurred at roughly the same rate for low and high lists. Second, the high-strength pairs in Experiment III were not necessarily the primary associates of the stimuli, so that a 'normative' guess would not match the actual response being called for. Third, and perhaps most convincing, is the fact that high-strength lists of approximately the same normative level have failed to produce superior retention.[22] If guessing is offered as an explanation of the present experiments, it cannot accommodate the lack of effect reported in the earlier study. For these reasons, then, these data are taken to be a relatively substantial verification of the role of the letter-sequence interference in the forgetting of letter pairs.

---

[22] Underwood (see n. 7 above).

EMMERT'S LAWS: SIZE CONSTANCY VS. OPTICAL GEOMETRY

DANIEL J. WEINTRAUB and GERALD T. GARDNER, *University of Michigan*

Abstract. Two relationships may be formulated concerning the size of projected afterimages: Emmert's law of retinal size—an extension of optical geometry—and Emmert's principle of perceived size—a principle linked to the classical issue of size constancy. The distinction between the relationships, the inadmissibility of null-match procedures for assessing perceived size, and related methodological issues and improvements are discussed. Two experiments investigating the dual formulation are described. Data from a distance-adjustment null-match method indicated that the law of retinal size held with precision for the individual observer. A falsified cue used to mislead observers did not affect null-match estimates but did affect a concomitant measure of perceived afterimage size, in accordance with the dual formulation.

The size of a projected afterimage increases with an increase in the projection distance and decreases with a decrease in the projection distance.[1] This is the gist of what Emmert wrote; it has come to be known as Emmert's law. What Emmert intended to convey to us by his statements is of historical importance, but his principle is too vague to merit the status of law. The principle does merit amending, however. There is a sufficient logical and experimental basis to make a case for two explicit Emmert's 'laws.' We follow the lead of Edwards and Boring, and differentiate between Emmert's principle of perceived size (linked to the issue of size constancy) and Emmert's law of retinal size (a straightforward extension of optical geometry).[2]

(I.) Emmert's *principle of perceived size* holds that the perceived size of a projected afterimage tends to covary with the

---

Received for publication August 11, 1969. Preparation of the article was aided by USPHS Research Career Award 1-K03-MH-35,253-01, from the National Institute of Mental Health to the first author. Support for the research was provided by the National Science Foundation, Grant GB-3110.

[1] E. Emmert, Grössenverhaltnisse der nachbilder, *Klin. Monatsbl. Augenheilkunde,* 19, 1881, 443–450.

[2] W. Edwards and E. G. Boring, What is Emmert's Law? this JOURNAL, 64, 1951, 416–422.

perceived distance of the surface onto which the afterimage is projected, both variables increasing or decreasing together—a relationship between perceived size and perceived distance similar to that produced by a physical stimulus which maintains a constant retinal size regardless of physical distance. A version of the size-distance invariance hypothesis states that a given retinal size (R) determines a unique ratio of perceived size (s) to perceived distance (d):

$$R = k(s/d). \qquad [1]$$

In order to clarify Emmert's principle of perceived size, the simple relationship of Equation 1 is used in the discussion that follows.

(II.) Emmert's *law of retinal size* holds that at the surface on which an afterimage is projected, the physical size (S) of a stimulus generating a retinal size (R) which matches the retinal size of the afterimage is directly proportional to the physical distance (D) of the projection surface. That is,

$$S_{\text{(equivalent physical size)}} = (R/K)D. \qquad [2]$$

(See Fig. 1, an instance of optical geometry applied to afterimages; since an afterimage has no physical size to 'cause' its retinal size, the physical size of the comparison has been labeled 'equivalent physical size' above.) Emmert's law of retinal size should always hold, regardless of the perceived size or perceived distance of the afterimage. Ultimately, however, two Emmert's relationships are necessary because the world of experience is not a perfect replica of the physical world. Afterimages further complicate matters of size and distance because an observer is concerned with the size, not of a physical stimulus, but of the aftereffects of a physical stimulus.

Three issues which permeate the literature on Emmert's 'law' can be pinpointed: What did Emmert really mean by it? What can be inferred about perceived size from various reported-size measures? Do the psychophysical data confirm Emmert's law of retinal size and Emmert's principle of perceived size? In response to the first issue, let us defer to the excellent discussions already provided by Edwards and Boring (see n. 2) and by Epstein, Park, and Casey.[3]

In response to the second issue, however, on inferring perceived size from a reported-size measure, let us start by supposing that a

---

[3] W. Epstein, J. Park, and A. Casey, The current status of the size-distance hypothesis, *Psychol. Bull.*, 58, 1961, 491–514.

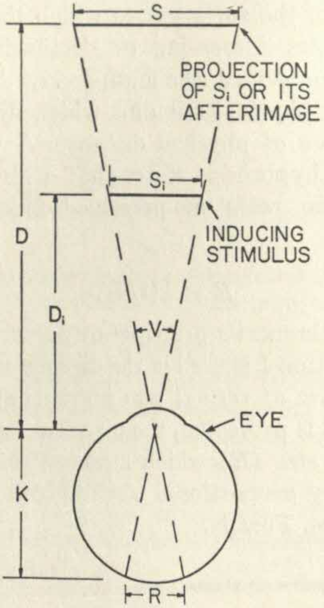


Fig. 1. Optical geometry

yardstick is superimposed upon an object and that the observer is asked to give the size of the object in inches. This is a form of null match; the observer counts the number of units subtended by the object. Or, he takes calipers, bracketing the object to make his null match, and the experimenter applies the yardstick to the calipers. Obviously, these techniques are those of basic linear measurement. Reported size will correspond very closely to physical size *regardless* of the physical distance between observer and object, and *regardless* of the perceived size or the perceived distance of the object. These techniques are relevant in testing Emmert's law of retinal size (see Young's bracketing method)[4] but not in testing Emmert's principle of perceived size. Having thus excluded null-match procedures (superposition and side-by-side matches) in assessing perceived size, we must further assume that the distance and size of a comparison stimulus are correctly perceived by the observer (perceived distance matches physical distance, perceived size matches physical size) if we are to use the physical size of the comparison stimulus as a measure of perceived size. This assumption of perfect size and distance constancy is perhaps never justifiable, but it may often be reasonable for nearby objects viewed in natural surroundings.

---

[4] F. A. Young, The projection of after-images and Emmert's law, *J. gen. Psychol.*, 39, 1948, 161–166.

In response to the third issue, on psychophysical evidence for the two Emmert's relationships, let us begin by noting that in 1948 Young published the first of a series of experiments assessing Emmert's law of retinal size by means of a superior experimental procedure (see n. 4). Young's observers bracketed the left and right edges of their afterimages by adjusting two pinpoint spotlights. This null-match adjusted-size procedure was employed on afterimages projected onto surfaces at different physical distances. Averaging across observers, Young presented the linear distance between the spotlights (equivalent physical size) as a function of physical distance to the projection surface and found that Equation 2 adequately described the averaged data. Several points were significantly different from their predicted values. But more disturbing was the finding that the data for each observer deviated a great deal from predicted values.

In contrast to Young's quantitative work, many other studies have been directed toward demonstrations which support a logical argument. Price carefully reviewed this literature and explicitly distinguished between the two possible interpretations of afterimage size: apparent size ('perceived size' in our terminology) and physical size ('equivalent physical size' is a preferred term since an afterimage has no physical counterpart); he proposed two corresponding laws.[5] In his experiments, distance cues were artificially manipulated so that the afterimage projection surface at the farther physical distance was at the closer perceived distance, and equivalent physical size was measured with a caliper method similar to Young's. Price did not present the resulting data but reported that the null-match reported-size measures (caliper technique) of the afterimages did not closely approximate the predicted values. However, the observers, using a rating scale, rated the afterimage at the closer perceived distance as smaller.

The experiments reported in the present paper applied an improved null-match technique in an effort to obtain precise corroborative evidence for Emmert's law of retinal size—evidence that neither Young nor Price had obtained. It should be possible to find experimental results that closely conform to the law of retinal size, a relationship which has always been considered valid on a priori grounds. The second experiment reported below attempted to demonstrate in a parametric study that two different reported-size measures, given sequentially on every trial, can produce disparate values for the same afterimage, each value presumed to correspond to the appropriate Emmert's relationship.

---

[5] G. R. Price, An Emmert's law of apparent size. *Psychol. Rec.*, 11, 1961, 145–151.

EXPERIMENT I

Instead of providing projection planes at different distances and requiring the observer (*O*) to report afterimage (AI) size at each of these distances, Experiment I utilized an alternative method (described by Gardner and Weintraub) in which different physical sizes were presented and *O* adjusted the projection distance appropriately.[6] A card on which a square was drawn was mounted on an apparatus that permitted *O* to control the distance of the card. After a square AI was induced, the *O* was asked to adjust the distance (*D*) until the square on the card (*S*) just surrounded his projected AI. With squares of different size, it was then possible to plot the function relating *D* to *S*.

Method

Subjects. The *O*s were 20 volunteers, men who were undergraduates at the University of Michigan.

Apparatus and procedure. The *O*'s head was immobilized by a frame that braced jaw and forehead. Below *O*'s eye level was a 49-in. track extending horizontally away, parallel to his line of sight. A channel in the track contained a sliding shoe bearing a vertical stimulus card. With a knob connected to a cord-and-pulley system, *O* could rapidly and accurately adjust the distance of the stimulus card. He could also alter background illumination, by using a variable voltage transformer to control the luminance of a shaded 100-w. bulb.

Comparison stimuli were drawn in india ink on white rectangular cards, 8 × 8½ in. Each card displayed the outline of a square and a centered fixation cross that was ¼ × ¼ in. The squares were of five sizes: 2, 3, 4, 5, and 6 in. on a side.

A positive AI was induced by flashing into the *O*'s eye a Honeywell No. 64B electronic photoflash masked to provide a stimulus that was a 1-in. square with a centered $\frac{1}{4} \times \frac{1}{4}$ in. fixation cross. The strobe light (manufacturer's specifications: 50 lumen sec/ft^2 at 4 ft.) was covered with a sheet of white paper (luminous transmittance 23%) to reduce luminance. One eye of the *O* was used at a time (the other was covered), alternating eyes every 5 min. and reeliciting the AI by flashing the strobe. The distance from strobe to cornea (6.06–6.44 in; mean 6.28 in.) varied slightly from *O* to *O* because of skull-shape differences.

---

[6] G. T. Gardner and D. J. Weintraub. Tracking afterimage shrinkage during fading, *Percept. & Psychophys.*, 3, 1968, 361–363.

After the strobe was flashed, it was removed from *O*'s line of sight. The *O* was asked to adjust the distance of the stimulus card until his square AI projected onto the card appeared to coincide with the square outlined on the card. If he could not center the AI in the square figure, he was instructed to adjust the distance so that his AI and the square on the card appeared to match in physical size. After 5 min. of practice, the main experiment was run, consisting of 30 track-setting trials with each stimulus card presented three times to each eye. The order of cards was independently randomized for each *O*.

Altering background illumination from time to time can improve the visibility of an AI and increase its duration. The *O*s were given practice in this procedure and were permitted to adjust the background illumination in any way they wished in order to facilitate their task.

Results

For each *O* the plot of adjusted physical distance (D) as a function of equivalent physical AI size (S) very closely approximated a straight line. Pearson correlation coefficients between size and distance were high, with no *O* falling below $r = .968$, and 13 of the 20 *O*s falling at $r = .990$ or higher. As the square of the r statistic reflects the amount of variance accounted for by a linear relationship, the high r's obtained indicated a high degree of linearity. This was further supported by an analysis of linear regression on the data for the *O* with the lowest correlation: deviation from linear regression was not significant ($F < 1.0$). Slopes of individual best-fit straight lines ranged from 5.74 to 7.26. A 95% confidence interval for the y-intercept of each *O* was established; in all cases, the interval around the obtained intercept included zero. The straight-line function for the combined data of all 20 *O*s had a slope of 6.53 and an intercept of .16 in., the 95% confidence interval for this intercept including zero.

Discussion

In principle, one could use physiological techniques to verify the match between the retinal size of an afterimage and of the image of a comparison stimulus, as required in a test of Emmert's law of retinal size. Null-match techniques like those used in the present experiment have *O* perform the psychophysical equivalent by lining up edges in his own visual field. In Equation 2 for a given afterimage, R/K can be considered a constant which is equal, by simple optical geometry, to the size of the inducing stimulus divided by its distance from the eye (see Fig. 1), $(R/K) = S_i/D_i$. A form of Emmert's law suitable for psychophysics is, then,

$$S_{(\text{equivalent physical size})} = (S_i/D_i)D\,. \qquad [3]$$

(Emmert's law of retinal size corresponds, strictly speaking, only to comparison stimuli, *i.e.* to a physical stimulus which at the given projection distance matches the afterimage.) Equation 3 contains the built-in assumption that at the distance of the inducing stimulus, the retinal sizes of afterimage and inducing stimulus are identical. Such is not necessarily the case. For example, if an afterimage is formed by staring at a bright light for many seconds, eye movements may smear the afterimage into a larger retinal size.[7] Brighter inducing stimuli may provide smaller afterimages.[8] The distance between inducing stimulus and the eye (more precisely, the center of the entrance pupil of the eye) must be measured carefully and the eye must remain fixed at that location during testing.

Because of the above limitations, a reasonable and usable form for Emmert's law of retinal size is:

$$S_{(\text{equivalent physical size})} = aD + b\,. \qquad [4]$$

In words, all size matches are presumed to lie on a straight line when plotted against D. A constant error in assessing eye-to-screen distance shows in the data as $b \neq 0$, and if $a \neq (S_i/D_i)$ then one may suspect either the measurement of S_i/D_i or the correspondence of the retinal size of the afterimage with the retinal size of the inducing stimulus (as in Young, see n. 8). However, Emmert's law of retinal size does not require this correspondence.

The results of Experiment I closely conform to the appropriate rearrangement of Equation 4 which makes D the dependent variable:

---

[7] But see F. A. Young, Studies of the projected after-image: I, Methodology and the influence of varying stimulation times, *J. gen. Psychol.*, 46, 1952, 73-86.

[8] F. A. Young, Studies of the projected after-image: III, Projection over large distances, *J. gen. Psychol.*, 47, 1952, 207-212.

$$D = (1/a)S - (b/a). \qquad [4a]$$

The obtained physical distances were directly proportional to equivalent physical AI size. Furthermore, this relationship held with considerable precision for each of the *O*s, as well as for their grouped data.

Although the obtained intercepts approximated zero, yielding a zero value for the constant b/a in Equation 4a, the obtained slopes $1/a$ were somewhat greater than would be predicted by $1/a = (D_i/S_i)$. This suggests that our efforts in measuring and controlling S_i and D_i were not sufficiently stringent (see below). Previous attempts to investigate Emmert's law of retinal size have assumed that confirmation requires agreement of each *O*'s data with Equation 3, yet the methodological details discussed above were not adequately controlled.

When the physical distance of the projection card is varied, afterimages phenomenally appear to expand and shrink, much to the surprise of some observers. The procedure of Experiment I may therefore seem to be assessing both perceived AI size and equivalent physical AI size. But as earlier remarks stressed, a null-match technique is pertinent only to equivalent physical size and is insensitive to any concomitant changes in perceived size. Regardless of what observers claimed to perceive, when the afterimage appeared to coincide with the outline of the square on the card, there can be little doubt that the retinal contour corresponding to the square on the card was superimposed upon the retinal border corresponding to the square afterimage. Perceived size and perceived distance are irrelevant to the outcome of such superimpositions, and Experiment II underscores the point.

EXPERIMENT II

Observers projected their afterimages onto cards placed in the apparatus of Experiment I. However, only the surface of the card was visible to a single eye, a reduction screen blocking the rest of the field. Mounted on each card was the replica of a postage stamp, half, normal, or double size. It was assumed that if the observer accepted the stamps as authentic, then judgments presumed to assess perceived afterimage size would be influenced by stamp size while null-match settings would remain unaffected.

Method

Subjects. Ten undergraduate and graduate students (men) at the University of Michigan served as *O*s.

Apparatus and procedure. The apparatus from Experiment I was used, with modifications and additions (see Fig. 2). The track assembly was offset to the left and was masked with a black cardboard shield. As the track was visible to *O* as he entered the room, a dummy extension was added to make it appear longer. He viewed the projection surface (stimulus card) with his left eye through a 6-in.-long black reduction tube, 2 in. in diameter. To the right of the track, he could have an unrestricted view with his right eye of a size-adjustment device placed at a distance of 80.75 in. This device consisted of two black metal vanes mounted (like curtains) on a traverse rod in front of a white background. Using a cord-and-pulley system, *O* could generate an equilateral diamond (a square tilted 45°) of any size up to 14 in. on a side, by changing the amount of overlap of the vanes and exposing more of the angular cutout in each vane.

The *O*'s right eye was first covered, and a diamond-shaped AI was induced in his left eye with the strobe unit at a distance of 6 in. The strobe was masked except for a diamond-shaped orifice (a 1-in. square tilted 45°). The strobe was removed from the line of sight, and *O* adjusted the distance of the stimulus card, started at a randomly determined position on the track, until AI size matched stimulus size as in Experiment I. When *O* was satisfied with

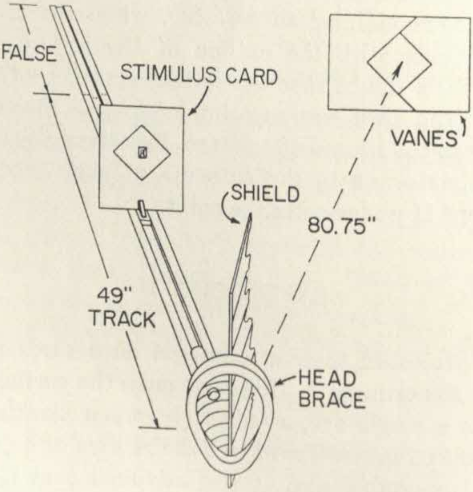


FIG. 2. Apparatus used in Experiment II

his setting, his left eye was covered, his right eye was uncovered, and he raised the room lights to the brightest setting. The *O* now adjusted the vanes to the perceived size of the AI as held in memory. After an initial adjustment, he was allowed to take one final look at the stimulus card; if he requested it, his right eye was covered and left eye uncovered for 5 sec., whereupon the eye coverings were reversed. Another AI was induced every 5 min.

The *E* was especially careful to explain the desired vane procedure; if *O* thought that the AI appeared to be 5 in. on a side at the surface of projection, he was to generate a vane opening such that it would be 5 in. when measured at the plane of the vanes. To emphasize the instructions, *O* was given three practice trials at generating a diamond which would be the same length on a side as a foot rule he had been shown at close range. With *O* looking on, *E* measured the opening each time with the rule and reported the actual setting to *O*. After this there were three practice trials of the entire experimental procedure using a special sample stimulus card.

Stimuli consisted of smooth-surface rectangular cards 15 × 17½ in., on each of which was the outline of a diamond (a square tilted 45°) drawn with india ink. Four sizes of diamonds were used: 2, 3, 4, and 5 in. on a side. At the center of each card was mounted a colored replica of a five-cent postage stamp. Each stamp size (half, normal, and double; all produced photographically) was used with each diamond size, making a total of 12 different stimulus cards. The thickness of the india-ink outline forming the diamonds was directly proportional to the size of stamp used regardless of diamond size. The sample card for the practice session bore a 3½-in. diamond and a normal-size stamp. This card was first shown to *O* in his full view and at his arms' length. The *E* explained that the stamp was put on the card to aid his judgments; no *O* expressed suspicion as to the authenticity of this stamp. The *O* saw the other stimulus cards only through the reduction tube, which did not reveal the edges of the cards.

Each of the 12 stimulus cards was used four times with each *O*, the order of presentation of cards independently randomized for each *O*. An experimental session lasted approximately 1½ hr.

Results

The data for the track (*i.e.* null-match) settings for each *O* as well as for the entire group showed the same precise, linear plot of physical distance and equivalent physical size of AI as was obtained in Experiment I. Furthermore, these relationships held independently of the size of the postage stamp at the center of the stimulus (see Fig. 3). A repeated-measures analysis of variance (Table I) indicated that the effect of diamond size was significant and accounted for over 96% of the variance, while the effect of postage-stamp size was not significant and did not account for a measurable proportion of the variance. In addition, the linear component of the effect of diamond size was significant, whereas the nonlinear com-

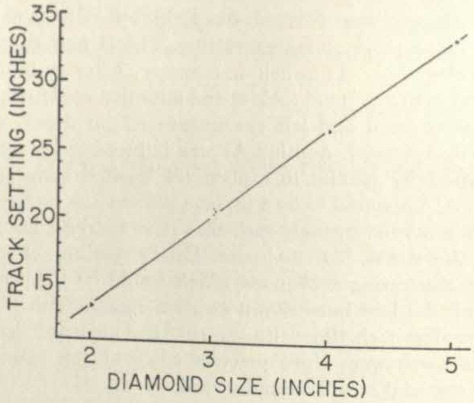


FIG. 3. Track-setting data of Experiment II. The triples of points show the data for each stamp replica. The three straight lines, one for each stamp-replica size, are so nearly coincident that only one line is shown.

ponent was not significant. The interaction of observers and diamond size was significant, showing that the plots for individual *O*s differed in slope.

The plot of the composite data for the vane settings as a function of diamond size is shown in Fig. 4. For any given stamp size, vane settings increased with physical stimulus size. More important,

TABLE I

SOURCES OF VARIANCE: EXPERIMENT II†

Source	*df*	Track settings		Vane settings	
		F	ω^2	*F*	ω^2
Between *O*s	9	15.36*	0.84%	163.13*	32.21%
Within *O*s	470				
St (Stamp Size)	2	0.08	0.00%	60.57*	29.70%
St × *O*	18	0.61	0.00%	8.37*	4.39%
Di (Diamond Size)	3	1,942.87*	96.08%	47.12*	17.20%
Linear	1	3,046.51*		56.32*	
Other	2	3.47		0.88	
Di × *O*s	27	2.11*	0.26%	4.70*	2.93%
St × Di	6	0.45	0.00%	2.02	0.60%
St × Di × *O*s	54	0.81	0.00%	2.44*	3.44%
Residual	360		2.82%		9.53%
Total	479				

* $p < .01$.

† Entries in ω^2 columns represent percentage of estimated true variance.

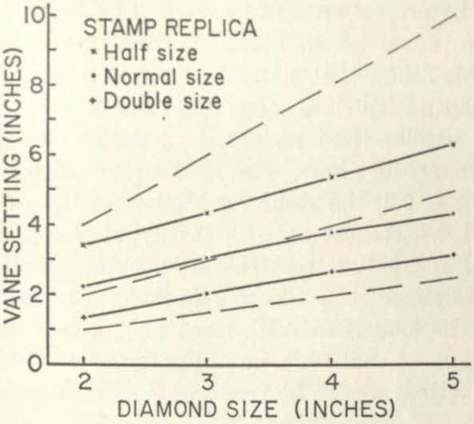


FIG. 4. Vane-setting data of Experiment II. The dotted lines represent predicted values.

however, the three separate functions that emerged indicated that vane settings increased as postage-stamp size decreased. Inspection of the data showed that this relationship held for individual *O*s, with only 8 means of a total of 120 showing misorderings. The results of a repeated-measures analysis of variance (Table I) indicated that the effect of diamond size was significant and accounted for over 17% of the variance; the linear component of this factor was the only significant trend. The effect of stamp size was also significant, accounting for almost 30% of the variance. The interaction of stamp size and diamond size was not significant, indicating that the differences between the slopes of the three lines in the vane plot were not significant.

Discussion

Our strategy had been to obtain two concomitant measures on each projected afterimage and to show that the measures were differentially but predictably influenced by the independent variables. For the first measure, the prediction was that the track settings (used also in Experiment I) should conform to Emmert's law of retinal size and that the data should therefore fit Equation 4a, showing a linear relationship between the physical size of the diamond on the projection surface (*S*) and the track settings necessary to make the afterimage appear to coincide with the diamond on the projection surface (D). The physical size of the postage stamp should be an irrelevant variable. Relevant data in Fig. 3 and Table I show that all predictions were upheld.

Young's quantitative data suggested that the retinal size of the afterimage is smaller than its predicted value (see nn. 4 and 8). Track settings reported here were in the same direction. Although $(1/a) = D_i/S_i = 6.3$, the slope for Experiment I was 6.5, and for Experiment II 6.7, consistent with the hypothesis that the retinal size of the AI is smaller than the retinal size associated with S_i. Afterimage shrinkage is a likely culprit for at least part of the discrepancy. The longer an AI is used for testing, the greater the shrinkage (see n. 6). All such interpretations of data assume Emmert's law of retinal size to be true and Equation 4a to be the suitable formula.

The second measure, the vane settings, was intended to assess perceived size. (Unfortunately there is no technique comparable to the precision of the null-match track settings with which to measure perceived size.) The vanes were less than 7 ft. away, in full view to each observer and with all the cues to size and distance usually present in a room. Therefore, as discussed above, it can be inferred that vane settings are an adequate indicator of perceived size. The dotted lines in Fig. 4 are based upon a simplified model of perceptual experience which assumes that the postage stamps were perceived by the observers as authentic and that the perceived distances and perceived sizes of the diamonds/afterimages followed in perfect accord. The data of Fig. 4 and Table I show that both stamp size and diamond size influenced vane settings; the results approximate the simplified model, although the interaction of stamp size with diamond size should have proved to be significant, which it was not.

Inferring the effect of stamp size on perceived AI size requires caution, however. The data do allow rejecting the hypothesis that stamp size was an irrelevant cue. This hypothesis would have been true if the observers had employed either of the two following strategies. (1) They might somehow have ascertained the true distance to the projection surface containing the stamp replicas (*e.g.* ineffective reduction conditions) and ignored stamp size; the data of Fig. 4 would then have collapsed to a single nonhorizontal line instead of the three functions obtained. (2) They might have made null matches with the vanes, as they did with the track settings. Though the AI was induced into the left eye and the vanes were observed by the right eye, both eyes contribute to a percept and an observer might still 'project' an AI to the surface of the vanes to make a null match. (Bringing the room lights to full illumination during a vane setting was intended to reduce the perceptibility of an afterimage.) If such null matches were made, then both diamond size and stamp size would have been irrelevant variables, with observers giving virtually the same vane setting each time, as predicted from the law of retinal size (a single horizontal line in Fig. 4 at a vane setting of 13.46 in.). On the other hand, it could be argued that the three functions in the data resulted from a strategy in which the observers estimated the length ratio between diamond and postage stamp ('how many postage stamps tall was the diamond?') and then tried to reproduce that ratio by adjusting the vanes relative to the remembered size of a postage stamp, a procedure which has features in common with null matches. The similarity seems only superficial considering the role of perceptual, memory, and cognitive factors not found in true null matches. Also, it is not clear how strategies of this kind would differ from a 'purely perceptual effect.'[9]

We feel that the observers accepted the stamp replicas as authentic and that they therefore experienced sizes and distances appropriate to a physical world which ordinarily would have provided such retinal size cues. In any case, Experiment II was not intended to submit Emmert's principle of perceived size to a critical test. The simplistic equation of the size-distance invariance hypothesis was selected primarily for illustrative purposes. (Vane-setting data would have to be supplemented with estimates of the perceived distance of the projection surface for a test of Emmert's principle.) The main thrust of the experiment was to show clearly and unambiguously that different measures of reported size, given concomitantly *to the same stimulus*, provide different data and that, therefore, two Emmerts' relationships are required—one tied very precisely to retinal size alone, the other tied only indirectly to retinal size. To underscore the two Emmert's relationships and their distinction, we conclude with the following counterintuitive example. A telescope shrinks afterimages! That is, if an AI is induced normally and then the observer projects his AI through a telescope, the equivalent physical size of the AI will shrink in proportion to the inverse of the magnifying power of the telescope (if the telescope does not alter the constant K). Perceived size will also tend to shrink, because a telescope compresses perceived distance. These results are in complete agreement with the two Emmert's relationships. The example also demonstrates why the term 'equivalent physical AI size' must replace the term 'physical AI size'; a telescope cannot change the retinal size of an AI already fixed on the retina.

---

[9] See W. Epstein, *Varieties of Perceptual Learning*, 1967, Ch. 2, for a discussion of the influence of familiar size on perceived size.

EFFECT OF STEREOSCOPIC CUES ON PERCEIVED WHITENESS

DONALD H. MERSHON and WALTER C. GOGEL,
University of California, Santa Barbara

Abstract. It has been found that the introduction of stereoscopic cues between an induction object and a test object will reduce the magnitude of the simultaneous whiteness contrast between the two objects. The present study both verified this result and demonstrated that the reduction in contrast could not be explained as the direct result of the small directional displacements that were necessary in order to produce the stereoscopic cue. The results from the present study also support the conclusion that the perception of depth, and not the stereoscopic cue per se, was the necessary condition for the change in the magnitude of the whiteness contrast.

If a large black disk is presented in a dark surround and is illuminated so that white light falls only on the disk, the disk will appear white or whitish gray. If a small object of greater luminance than the large disk is placed on the disk, the apparent whiteness of the disk decreases. This darkening of the large disk in the presence of the small object is known as the Gelb effect. The Gelb effect was long regarded as evidence for the view that knowledge of the conditions of illumination is important in the perception of whiteness, with the insertion of the small object into the beam of light providing the information necessary for this perception. Thus, the Gelb effect was usually considered to be an example of whiteness constancy and was believed to be dependent upon central rather than peripheral (retinal) factors. Stewart, however, demonstrated that the Gelb effect follows the laws of simultaneous contrast (a phenomenon usually believed to take place at a peripheral level).[1] Two factors important in simultaneous contrast are the frontoparallel (directional) separation of the test object from the induction object and the size of the induction object. Stewart

---

Received for publication November 3, 1969. The study was supported by Grant 1-R01-MH15651-02 from the U. S. Public Health Service. The authors wish to thank John Brendel for his help in collecting the data.

[1] E. C. Stewart, The Gelb effect, *J. exp. Psychol.*, 57, 1959, 235–242.

showed that these factors were also important variables in the Gelb situation. As expected of a contrast phenomenon, the Gelb effect decreased as the small induction object was moved directionally from the center of the large disk and also as the size of the induction object decreased. It may be concluded, therefore, that the Gelb effect is an instance of the operation of simultaneous contrast.

A recent study by Gogel and Mershon demonstrated that the magnitude of the Gelb effect decreases if the stereoscopic cue is used to separate the test object and induction object in perceived depth.[2] This finding supports the view that whiteness contrast, as exemplified by the Gelb effect, can be modified by central processes of neural organization at or beyond the level of organization responsible for the stereopsis.

Whenever a cue to depth is varied to produce variations in perceived depth, some change in the stimulus conditions must occur. It is essential, therefore, in studying the effect of a separation in depth on the Gelb effect, to determine that the change in the Gelb effect is due to the generation of perceived depth and is not due to the change in the stimulus conditions per se. For example, the use of the stereoscopic cue to make the induction object appear at a different distance than the Gelb disk requires a lateral displacement of the induction object (in opposite directions in each eye) with regard to the Gelb disk. In the study by Gogel and Mershon (p. 16), it was argued that this small displacement (a maximum of 17 min. of arc in each eye) could not account for the obtained results. The present study directly tested this factor by comparing the results from monocular and binocular observation of the objects under the same conditions of physical displacement. It was expected that under monocular conditions, the induction object would appear at approximately the same distance as the Gelb disk due to the action of the equidistance tendency.[3] The lateral displacement of 17′ would still be present, however, in the observing eye. If the results of monocular observation were significantly different from those of stereoscopically disparate binocular observation, then the amount of lateral displacement present in the disparate binocular condition could not have produced the effect

[2] W. C. Gogel and D. H. Mershon, Depth adjacency and simultaneous contrast, *Percept. & Psychophys.*, 5, 1969, 13–17.

[3] The equidistance tendency is the tendency for objects to appear equidistant in the absence of effective cues to depth; see W. C. Gogel, Equidistance tendency and its consequences, *Psychol. Bull.*, 64, 1965, 153–163.

which has been attributed to the stereoscopic separation. To investigate the generality of the results obtained by Gogel and Mershon, a square rather than a small disk was used as the induction object in the present study.

An additional question which this paper sought to answer concerns the choice of the most appropriate measure for comparing the magnitude of the Gelb effect under different conditions. Since the basic description of the phenomenon involves the *darkening* of the disk when the small object is introduced into the field, it would seem reasonable that the appropriate measure would be one which involves a specification of this change. That is, the difference between the perceived whiteness of the disk alone and of the disk under a condition of contrast should be compared across conditions. This was the procedure followed by Gogel and Mershon. This may not, however, be the most sensitive measure for comparing the effects of separations in depth upon contrast. Stewart, for example (see n. 1), directly compared the whiteness matches obtained under the various conditions of lateral displacement of the induction object. This procedure would be a better measure to use if there were factors determining the perceived whiteness of the disk in isolation which did not also determine the perceived whiteness of the disk under the conditions of induction. At least one such factor seems possible. Newson has pointed out that in isolation objects varying in reflectance from 18% to 45% all appear as "reasonably good" whites (*i.e.* *O*s match them by Munsell values ≥ 8.0).[4] The data of the present study and of the study by Gogel and Mershon were examined, therefore, for such 'extraneous' factors in the perception of isolated objects.

METHOD

Apparatus. The apparatus used in the present study, and illustrated in Fig. 1, was basically the same as that used by Gogel and Mershon. Essentially, this consisted of a dimly luminous disk of light (A_1, .79 ftL.) with a diameter of 20.5 cm. (4° 1′) at a constant distance of 284 cm. from *O*, and a much brighter square (A_2, 156 ftL.) at a distance of either 284 cm. or 197 cm. from *O*.

---

[4] L. J. Newson, Some principles governing changes in the apparent lightness of test surfaces isolated from their normal backgrounds, *Quart J. exp. Psychol.*, 10, 1958, 82–95. The Munsell values usually associated with a range of reflectances of 18%–45% is approximately N 4.8/ to N 7.1/, when the samples involved are presented against a uniform background.

The test disk was produced by an aperture (A_1) located in front of a homogeneous black paper uniformly illuminated by a hidden fluorescent light (L_1). The induction square was produced by a movable aperture (A_2) located in front of a luminous surface (L_2) which remained at a constant distance from *O*. The surfaces used for the test disk and induction square were homogeneous and were perceived as localized at the distance of the apertures A_1 and A_2 respectively.[5] The *E* could position A_2 at either of two distances from *O*. The physical size of the square was different at the two different distances so as to subtend a constant visual angle (1° 23′ on a side). The square was directionally superimposed on the disk by means of a partially transmitting, partially reflecting mirror (M) that permitted *O* to see the square and disk simultaneously. A schematic front-view drawing of the square and disk, as seen from the position of *O*, is shown in the inset of Fig. 1. Throughout the experiment nothing was visible to *O* except the square and disk; the remainder of the field was totally dark. Shutters were arranged so that *O* could view (either binocularly or monocularly) the disk alone or the square and disk simultaneously. When the square was at the same distance as the disk, it was directionally centered on the disk in each eye. When the square was in front of the disk (by 87 cm.), the center of the square was physically displaced to the right or left of the center of the disk by 17′ in both the left eye and the right eye respectively, generating the binocular disparity cue between the two objects. The darkened observation booth contained a head-and-chin rest and a pair of artificial pupils (1.0 mm. in diameter), with the lateral separation between the artificial pupils adjustable to the interpupillary distance of each *O*.

Observers. The 45 *O*s used in the experiment were college students with a stereoacuity of at least 40 sec. of arc and a vernier acuity of at least 20/20 (both near and far) in each eye as measured with a Keystone orthoscope. The

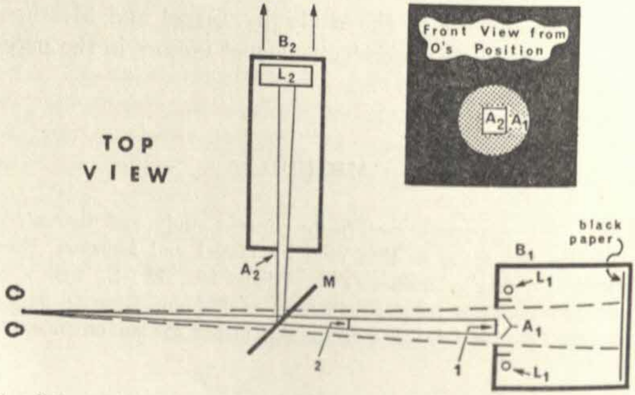


FIG. 1. Schematic top-view drawing illustrating apparatus and conditions

interpupillary distance of each O was measured with a Mark I Naval Interpupillometer.

---

[5] W. Lichten and S. Lurie, A new technique for the study of perceived size, this JOURNAL, 63, 1950, 280–282.

Procedure. Before being presented the disk or square, each O was taught a scale of whiteness, using six Munsell patches along the black-to-white dimension.[6] The blackest patch was numbered 1 and the whitest 6, with the other patches numbered appropriately. After the training period, which took place in a separate room, O was taken to the observation booth where he received instructions about his task. He was then left in the dark for 5 min. before being presented with the first condition of the experiment. For each condition of the experiment, the disk was presented alone first, followed after 30 sec. by the simultaneous presentation of the square and disk. The judgments of whiteness were always made with respect to the disk, never with respect to the square. When the disk was presented alone, O indicated its apparent whiteness (W') using the previously learned scale. When the disk and square were presented simultaneously, in addition to the W' of the disk, O indicated verbally the perceived distance (D') of the disk from himself and the perceived depth (d') between the square and disk.

Three conditions were used. Each condition consisted of the presentation of the disk followed by the simultaneous presentation of the disk and square. In the condition labeled *Binocular-Same,* the disk and disk-square sequence was presented binocularly with the square and disk physically at the same distance from O (Position 1 in Fig. 1). In the *Binocular-Different* condition, the disk and disk-square sequence was presented binocularly with the square 87 cm. *nearer* to O than the disk. In the *Monocular* condition, the disk and disk-square sequence was presented monocularly to the right eye, with the square physically 87 cm. in front of the disk. Following each disk and disk-square sequence, the O was in complete darkness for 2½ min. before being presented with the next condition.

With 15 Os (Group I) the order of presentation of conditions was Binocular-Different, Binocular-Same, and Monocular. With 15 other Os (Group II) the order of presentation was Binocular-Same, Binocular-Different, and Monocular. With the remaining 15 Os (Group III) the order of presentation was Binocular-Different, Monocular, and Monocular. The Os were run individually in randomly sequenced blocks, with one O from each group in each block. Note that the monocular presentations always followed the binocular presentations. If this had not been the case, the monocular presentations could have affected the subsequent judgments of W' made under binocular conditions, by producing differential states of dark adaptation in the two eyes. With the third order of presentation listed above, the Monocular condition was presented twice in succession in order to evaluate the possible effect of the order of presentation upon changes in W' within the same condition.

RESULTS

The means and standard deviations of the W', d', and D' results from two of the three orders of presentation of the different conditions are given in Table I. The results from the remaining order of presentation are shown separately in Table II. In the tables the columns labeled Diff. are the average differences between the W' values shown in the columns labeled Square Absent and Square Present, and the standard deviations of these differences. For example, in Table I, the average decrease (darkening) in the perceived whiteness of the disk as a consequence of adding the square was .87 units for the first binocular presentation of the disk and disk-square sequence when the square was physically 87 cm. nearer to *O* than the disk.

---

[6] See Gogel and Mershon (n. 2), 14.

Analyses of the results of Table I. The d' results in Table I indicate that the stereoscopic cue was effective in the Binocular-Different condition in producing a perception of depth between the square and disk. A positive d' indicates that the *O* perceived the square in front of the disk and a negative d' indicates the reverse. In both the Binocular-Same and Monocular conditions, the square, on the average, was perceived at approximately the same distance as the disk. This can be attributed to the zero stereoscopic difference in the Binocular-Same condition and to the equidistance tendency in the Monocular condition. It has been suggested, however, that the perceived separation of the two objects, without regard to the direction of the separation, is the significant factor in the adjacency effect on contrast.[7] If the average values of d' are computed without regard to sign, the results of the three conditions were Binocular-Different, 22.1 in.; Binocular-Same, 2.4 in.; Monocular, 9.0 in. These results are significantly different at the .002 level using the analysis of variance (treatments × subjects design, $F = 19.9$, $df = 2/58$).[8]

Similar analyses of variance were employed to determine the effects of the main conditions on both possible measures of whiteness contrast. Generally, the results for both measures were parallel. The overall average decreases in perceived whiteness of the Gelb disk (Table I) as a consequence of adding the square were Binocular-Different, .82 units; Binocular-Same, 1.36 units; Monocular, 1.18 units. These results differed significantly at the .002 level (two-tailed test, $F = 8.04$, $df = 2/58$). The overall average values of perceived whiteness of the Gelb disk (Table I) under the

---

[7] See Gogel and Mershon (n. 2), 14–15.

[8] E. F. Lindquist, *Design and Analysis of Experiments in Psychology and Education,* 1953, 156–166.

TABLE I

MEANS AND *SD*s OF PERCEIVED WHITENESS (W') OF DISK, PERCEIVED DEPTH (d') BETWEEN DISK AND SQUARE, AND PERCEIVED DISTANCE (D') OF DISK FROM OBSERVER

	Binocular—Different condition			Binocular—Same condition			Monocular condition		
	First presentation (Group I)			First presentation (Group II)			Third presentation (Group II)		
	Square absent	Square present	Diff.	Square absent	Square present	Diff.	Square absent	Square present	Diff
Mean W'	4.00	3.13	.87	4.05	2.42	1.63	3.72	2.42	1.30
SD of W'	.71	.52	.66	.55	.62	.62	.62	.60	.75
Mean d'		20.93			−1.93			−3.77	
SD of d'		16.67			3.41			7.65	
Mean D'		48.00			58.00			42.07	
SD of D'		33.95			34.13			27.44	
	Second presentation (Group II)			Second presentation (Group I)			Third presentation (Group I)		
	Square absent	Square present	Diff.	Square absent	Square present	Diff.	Square absent	Square present	Diff
Mean W'	3.61	2.83	.78	3.83	2.75	1.08	3.80	2.73	1.07
SD of W'	.52	.77	.78	.54	.76	.91	.70	.83	.93
Mean d'		23.20			2.20			−4.80	
SD of d'		18.33			5.47			19.67	
Mean D'		53.20			49.20			41.33	
SD of D'		32.04			27.99			22.93	

Note: On scale of W', full white is 6 and black is 1; d' and D' in cm.

TABLE II

Means and *SD*s of Perceived Whiteness (W') of Disk, Perceived Depth (d') between Disk and Square, and Perceived Distance (D') of Disk from Observer

	Binocular—Different condition			Monocular condition			Monocular condition		
	First presentation (Group III)			Second presentation (Group III)			Third presentation (Group III)		
	Square absent	Square present	Diff.	Square absent	Square present	Diff.	Square absent	Square present	Diff.
Mean W'	4.07	2.97	1.10	3.76	2.95	.81	3.68	2.71	.97
SD of W'	.44	.88	.95	.65	1.20	1.07	.67	.94	.94
Mean d' in cm.		11.57			−11.18			−14.03	
SD of d'		17.92			26.58			37.22	
Mean D' in cm.		38.93			28.87			27.27	
SD of D'		29.14			30.52			26.54	

Note: On scale of W', full white is 6 and black is 1; d' and D' in cm.

conditions of induction (columns labeled Square Present) were Binocular-Different, 2.98 units; Binocular-Same, 2.58 units; Monocular, 2.58 units. These results also differed significantly at the .002 level (two-tailed test, $F = 10.2$, $df = 2/58$). The Newman-Keuls procedure for making ordered comparisons showed the same patterns of significant differences in both analyses. The Binocular-Same and Monocular conditions were both significantly greater at (at least) the .05 level than the Binocular-Different condition, but in neither case was there a significant difference between the Binocular-Same and Monocular conditions.

Comparing the 'Diff.' scores of the first vs. the second presentations with the effects of stereoscopic separation counterbalanced (and thus eliminated), there appears to have been some effect of order: 1.25 units for the first vs. .93 units for second. Although this difference was not significant ($F = 1.66$, $df = 1/29$), it was in the direction expected from the results of Gogel and Mershon. There was clearly no effect of presentation order in terms of the 'Square Present' data: 2.78 units for first vs. 2.79 units for second.

The above results permit three conclusions. (1) The main effect of the Gogel and Mershon study, demonstrating a decrease in the Gelb effect as a consequence of the stereoscopic separation of the test object and induction object, is supported. (2) The reduction in the Gelb effect as a consequence of the stereoscopic cue cannot be accounted for by the slight directional displacement required to produce the stereopsis. (3) Both of the methods of measuring the Gelb effect seem satisfactory as far as the present study is concerned.

Analyses of the results of Table II. The results of Table II are, in general, more variable than those of Table I. It will be noted that the mean d' values had approximately the same *magnitude* for all the conditions. If the average values of d' are computed without regard to sign, the results of the three conditions were clearly similar: Binocular-Same, 14.8 in.; Monocular (second presentation), 15.6 in.; Monocular (third presentation), 17.6 in. If, as has been suggested, the perceived separation of the square and disk, rather than the stereoscopic cue per se, is the significant variable in determining the magnitude of the Gelb effect, it follows that the degree of whiteness contrast should have been approximately the same for these three conditions. In agreement with this interpretation, the analyses of variance did not show any significant differences among the conditions for either measure of contrast. More detailed comparisons (for example, between the two monocular conditions) were not warranted in light of this lack of significance. It should be noted, however, that the results of Table II are consistent with the notion that perceived rather than stereoscopic depth is the important factor in determining the main effect of this study.

DISCUSSION

The present study confirmed the view that stereoscopic separation is a significant factor in simultaneous whiteness contrast. From the present data it is difficult to see how the evidence for this conclusion can be attributed to any experimental artifact involved in the production of the depth separation by the steresocopic cue. Also, the data are consistent with the notion that it is the perceived depth between the test object and induction object, rather than the depth cue per se, that is significant in determining the differences in *W′*. More evidence is needed, however, before this latter conclusion can be accepted with confidence. The remainder of the discussion centers on (*a*) a consideration of the most appropriate measure to be used in reporting contrast in the Gelb situation and (*b*) a brief review of some other studies relevant to the same or similar phenomena.

Considerations of measurement. In the present experiment the results were basically the same regardless of which of the two measures of whiteness contrast was used. There was, however, a tendency (nonsignificant) for the 'Diff.' scores of second presentations to be somewhat less than the 'Diff.' scores of first presentations. This tendency was not present when considering only the data from the conditions of induction ('Square Present'). Since Gogel and Mershon found a significant effect of presentation order (first vs. second vs. third) using 'Diff.' scores, it is of interest to determine whether the same effects of order are present in the data of that study when the *W′* data of the conditions involving the induction object are compared directly. The results of this comparison are: first presentation, 2.94 units; second presentation, 2.95 units; third presentation, 3.01 units. These values are *not* significantly different by the analysis of variance (treatments × subjects design, $F = .483$, $df = 2/118$). Thus, the Gogel and

Mershon experiment found a significant effect of order of presentation on whiteness contrast only if the 'Diff.' data is considered. This result, in combination with the tendency found in the present study, suggests that some factor associated with order of presentation was present in both studies to affect the perceived whiteness of the Gelb disk when presented alone, but not when presented simultaneously with an induction object. The effect of this factor seems to be one of preventing the isolated disk from regaining its initial appearance of whiteness following a condition of contrast. If this analysis is correct, then the average perceived whiteness of the test disk when the induction object is present is a better measure for comparing the effects of the conditions of apparent depth separation than is the difference score. When the former measure is used to compare the effects of the three conditions of binocular separation of the Gogel and Mershon study (see their Table I, all presentations, columns labeled Small Disc Present), the results are significant at the .002 level (two-tailed test, $F = 8.34$, $df = 2/118$).

It is not possible, of course, from either the Gogel and Mershon study or the present one, to identify the factor that is affecting the perceived whiteness of the isolated object as a function of the number of presentations. One possibility lies in the fact that in both studies, 5 min. of dark adaptation preceded the first presentation, whereas only 2½ min. of dark adaptation preceded subsequent presentations. What is significant is that the factor involved did not operate directly upon the perceived whiteness of the disk under the conditions of induction. The importance of this finding is clear for the results of the present study, in which it was necessary always to present the Monocular condition following the binocular ones. Using the 'Square Present' data, the Monocular and binocular conditions can be compared without difficulties of interpretation arising from the unbalanced orders of presentation. Such comparisons show that, for Table I at least, the Monocular condition was clearly more similar to the Binocular-Same condition than to the Binocular-Different condition.

Related studies. Although the present work developed from an interest in extending the applications of the adjacency principle,[9] it is important to note that there have been earlier references to similar phenomena. Koffka, for example, discussed the question of the critical configurational variables in the perception of whiteness in terms of 'appurtenance,' *i.e.* the Gestalt notion of "belongingness."[10] Basically, he believed that the relationships most important in determining the perceived whiteness of an object are the relationships between the object and those other objects which most "belong with it." He remarks: "Clearly, two parts at the same apparent distance will, *ceteris paribus,* belong more closely together than field parts organized in different planes" (p. 246). Although it may be true that more adjacent objects 'belong' together more than less adjacent objects, it is believed that the adjacency principle offers a more definitive explanation of the phenomenon than the concept of 'appurtenance,' in that the adjacency principle is tied more readily to measurable perceptual variables.

---

[9] W. C. Gogel, Size cues and the adjacency principle, *J. exp. Psychol.*, 70, 1965, 289–293.

Wolff was concerned with showing that under certain conditions with a three-dimensional stimulus configuration, no effects of contrast would occur.[11] With both achromatic and chromatic stimuli, he found that test patches were not influenced by their backgrounds when the test patches were attached to stands set 105 cm. in front of the backgrounds. When the test patches were suspended from thin threads rather than being attached to the stands, he noted that the test patches were sometimes localized against the background and at other times in front of the background. Effects which may be attributed to contrast occurred only *when the test patches appeared in the plane of the background* (pp. 95–96). Unfortunately, Wolff presented no quantitative data on the perceived separations of his test patches and backgrounds. From his description of the apparatus, however, it can be calculated that the physical separation resulted in a binocular disparity of approximately 1° 20′. This is above the limits for fusion. It is difficult to evaluate the effect upon perceived whiteness of the resulting double images and large lateral displacements of the images in each eye.

Recently, Coren has concluded that a surface will be more affected by contrast from its surround when the surface is perceived as a figure and the surround as background, than when the reverse occurs.[12] To ensure that a particular portion of his configuration was perceived as figure, Coren sometimes used a steroscopic cue to displace the portion in depth from the background. Unfortunately, Coren did not measure perceived separation. Unless 'figureness' differs quantitatively as a function of the magnitude of stereoscopic separation, it is difficult to see how Coren's formulation would predict the present graduated results. Also, in both the current study and that of Gogel and Mershon, the object which was observed to undergo the effect of contrast was the *farther* and *larger object, i.e.* the object which should have been generally seen as the background.

---

[10] K. Koffka, *Principles of Gestalt Psychology,* 1963, 246–247, 260.

[11] W. Wolff, Ueber die kontrasterregende Wirkung der transformierton Farben, *Psychol. Forsch.,* 18, 1933, 90–97.

[12] S. Coren, Brightness contrast as a function of figure-ground relations, *J. exp. Psychol.,* 80, 1969, 517–524.

AGREEMENT AMONG SUBJECTS IN CHOOSING TRIGRAM LABELS FOR RANDOM SHAPES

BERNHARD KEMPLER, *Georgia State University*
and HERBERT KAYLE, *Emory University*

Abstract. Fifty-three undergraduates chose 'names' for 10 unfamiliar shapes from a list of five CVCs. Both the shapes and the CVCs were constructed and selected randomly. Each shape was singly presented 10 times for labeling. Results showed that a majority of the subjects were consistent (nonrandom) in selecting labels for the shapes and that there was significant between-subject agreement on how seven of the shapes should be named. The shapes were seen as opposites on a dimension and two of the CVCs were used as appropriate labels for the poles of this dimension. A representational theory of reference, based on perceptual similarities between label and referent, is proposed as a possible explanation of the results.

Members of the same language community show considerable agreement in choosing the appropriate connotation for a nonsense syllable from such dimensional opposites as 'large-small,' 'dark-bright,' 'hot-cold,' and so forth.[1] This has been interpreted as support for the hypothesis of phonetic symbolism—the hypothesis that there is "an intrinsic correspondence between sounds and meanings."[2] Whether such correspondence is based on broad language-dependent associations, on correlations among sense data actually present in the world, or on individuals' ability to discover perceptual similarity between vocal sounds and various attributes of referent objects, remains an unresolved issue.[3] Some investiga-

---

Received for publication February 17, 1969. The study was supported in part by funds from the Educational Research Fund (USOE), Emory University Research Fund, and NIH Grants MH-11210 and HD-02458. Early phases of the study were completed while the first author was also at Emory. The authors thank Mr. Alan Sockloff for his invaluable assistance with the statistical analyses.

[1] R. Brown, *Words and Things*, 1958; I. K. Taylor and M. M. Taylor, Another look at phonetic symbolism, *Psychol. Bull.*, 64, 1965, 413–427.

[2] I. K. Taylor, Phonetic symbolism re-examined, *Psychol. Bull.*, 60, 1963, 200–209.

[3] These suggestions were made, respectively, by Taylor and Taylor (see n. 1 above), by Brown (see n. 1), and by H. Werner and B. Kaplan, *Symbol Formation*, 1963.

tors have, however, argued that the demonstration of phonetic symbolism may itself depend on the fact that the subjects were provided with a judgment dimension.[4] Others have held that a meaningful referent (*e.g.* 'large *table*–small *table*') may also be required for agreement among subjects.[5] Consistent with either set of assumptions has been the practice of requiring subjects to match pairs of items, one contrasted phonetically, the other contrasted on some sensible dimension expressed as a pair of familiar antonyms like those above.

Could consistent sound-referent matching be obtained under other and less restrictive conditions? More specifically, could subjects agree in their selection of conventionally meaningless sounds as names for novel and unfamiliar referents, in the absence of a particular judgment dimension? In the present study, undergraduates labeled random shapes with 'names' from a randomly selected finite list of nonsense trigrams. Instead of matching contrasted pairs of items, they selected one of five possible names for 10 different shapes as each of these shapes was separately presented (10 times) for labeling. The main question was whether within- and between-subject agreement would indeed occur when (1) no bipolar judgment dimension was provided or intentionally represented in the stimuli, (2) referent terms were presented one at a time rather than by the traditional paired-comparison procedure, and (3) neither the referents nor the vocal sounds had a conventional meaning.

METHOD

Stimuli. The stimuli were 10 random shapes and five CVC trigrams. The *random shapes* were solid black polygons with straight sides, constructed by Attneave's Method I and using an 80 × 80 matrix and 12 initial points.[6] They were then photographed on Kodachrome film and presented on standard 35-mm. transparent slides. The *CVCs* were generated by successive random drawings of a consonant, a vowel, and a consonant, eliminating those combinations which were actual or phonetically real words in English. The five CVCs used in the study were HAB, SUD, BIZ, ZEQ, and TUS.

---

[4] *E.g.* R. Brown and R. Nuttall, Method in phonetic symbolism experiments, *J. abnorm. soc. Psychol.*, 3, 1959, 441–445.

[5] *E.g.* J. H. Weiss, The role of stimulus meaningfulness in the phonetic symbolism response, *J. genet. Psychol.*, 70, 1964, 255–263.

[6] F. Attneave and M. D. Arnoult, The quantitative study of shape and pattern perception, *Psychol. Bull.*, 53, 1956, 452–471.

Subjects. The *S*s were 87 men and women enrolled in three different undergraduate psychology courses. The experiment was administered separately to each class.

Procedure. The *S*s were seated in rows facing a projection screen at the front of the room. Each was given an answer booklet consisting of 100 items, each item being a random sequence of the five CVCs typed in capital letters. A cardboard template with a window was placed over the page to cover all items except the one on which *S* was working. The purpose of the template was to make *S*'s response on each item as independent as possible of every other and to guard against the possibility of his copying from other *S*s.

The 10 random shapes were projected one at a time at a constant rate on the projection screen by a Kodak 800 Carousel projector. Each shape was presented 10 times in random order for a total of 100 presentations, with the restrictions that the same shape never be presented twice in succession and that each shape appear five times in the first half and five times in the second half of the series. Each presentation of each slide was for 15 sec., 8 sec., or 5 sec. respectively for the three groups of *S*s. Before the first slide was presented, *E* verbally gave these instructions:

> "This is a study of the process of naming. You have each been given an answer booklet and a cardboard template. Place the template over the booklet so that only item number 1 is visible. Listen now and I will pronounce the five syllables for you. [*E* then twice pronounced the five syllables slowly, with the following pronounciation: HAB—haeb; SUD—s∧d; BIZ—biz; ZEQ—zek; TUS—t∧s.] There are 100 items in this booklet, each consisting of these five syllables.
>
> "On this screen I will show you, one at a time, 100 unfamiliar shapes. Your task is to select an appropriate name for each shape from the five syllables in your answer booklet. I would like you to adopt the attitude that this is a meaningful task and that you are assigning unfamiliar names to unfamiliar objects in such a way that the names seem right and appropriate to you. The shapes will be shown at a constant rate. Choose a name for each shape even when you feel unsure. Don't skip any."

On completion of the task, *S*s were asked to write down as best they could a brief account of the basis for their choices.

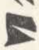
RESULTS

It soon became clear that a number of protocols could not be used because the *S*s had got 'out of step' on the shapes and the answer-sheet items: when the last figure was presented, these *S*s had either used up all their response items or had more than one left over. Since there was no way of determining where *S* had skipped, the protocols were considered invalid. The protocols for *S*s who had stayed 'in step' but had failed to make a labeling choice for one or more items were also thrown out. This resulted in a total of 53 valid protocols. Since many invalidated protocols occurred in the group that got the 5-sec. presentations, apparently because of time pressure, it was decided to ignore the duration of presentation and to collapse the three groups in the data analysis.

Given 10 trials for each shape and five trigrams to choose from, the binomial probability of selecting one of the trigrams six or more times is .006. Using this probability level as the criterion for consistent (*i.e.* nonrandom) labeling, the number of *S*s reaching this criterion for each shape was computed. These data are presented in Table I, which shows that Shape I, for example, was consistently labeled BIZ by 1 *S*, HAB by 25 *S*s, SUD by 10 *S*s, TUS by 5 *S*s, and ZEQ by 5 *S*s, for a total of 46 *S*s (of 53) who consistently chose one of the trigrams as a label for Shape I. On the other hand, only 20 *S*s found appropriate labels for Shape VI. The mean number of *S*s per shape showing consistent labeling was 34 of the total sample of 53. Using the somewhat less stringent criterion of selecting the same trigram five or more times (a probability of .03), the number of *S*s showing consistent labeling increases to a mean of 41 per shape.

TABLE I

Number of Subjects Selecting Trigrams as Labels on Six or More of Ten Presentations of each Shape

Shape	BIZ	HAB	SUD	TUS	ZEQ	*N*	χ^2	*p*
I	1	25	10	5	5	46	38.6	.01
II	6	1	4	7	13	31	8.5	N.S.
III	9	1	0	4	19	33	25.2	.01
IV	4	3	4	3	20	34	19.3	.01
V	5	11	8	5	2	31	9.4	N.S.
VI	4	1	2	6	7	20	5.3	N.S.
VII	4	11	9	4	1	29	13.5	.01
VIII	5	3	4	9	17	38	10.7	.05
IX	4	1	2	6	24	37	31.5	.01
X 	4	15	9	5	5	37	12.1	.05

Note: $N = 53$; the probability of selecting the same trigram on six or more of 10 presentations is $< .006$.

To determine the degree of agreement among the *S*s, chi squares were computed separately for each of the shapes.[7] These results are also shown in Table I. For example, a significant ($p < .01$) number of *S*s agreed that Shape VII should be called a ZEQ and that Shape I should be called a HAB.

Inspection of the data in Table I suggests that the seven shapes showing significant agreement were labeled either HAB or ZEQ by the largest number of *S*s. Furthermore, there was a noticeable trend for these two trigrams to be in a trading relationship with each other: shapes labeled ZEQ (III, IV, VIII, IX) tended not to be labeled HAB, and shapes labeled HAB (I, VII, X) tended not to be labeled ZEQ. These trends suggest that *S*s polarized at least seven of the shapes on some dimension and were able to discover in the trigrams HAB and ZEQ an appropriate connotative 'fit' with the poles of this dimension.

Two informal follow-up studies were performed (1) to ascertain whether the random shapes did contain a perceptual basis for bipolarization and (2) to obtain a description of the 'fit' between HAB and ZEQ and the perceptual dimension. In the first study, the seven shapes which had produced significant between-subject agreement were presented to 10 new *S*s with the instructions: "Put these shapes in two piles; put those together that belong together." In every instance *S*s placed the HABs (Shapes I, VII, and X) in one pile and the ZEQs (Shapes III, IV, VIII, and IX) in the other.

In the second study, the three HABs and three ZEQs (Shape VIII having been eliminated from the ZEQs to avoid possible effects of unequal numbers) were presented on mimeographed sheets; the HABs were placed together in one box and the ZEQs in another. The trigrams HAB and ZEQ were typed, one below the other, in the center of the sheet, followed by these instructions:

> "Here are two sets of shapes and two possible names for them. Pronounce the names to yourself first. Then choose the name that you feel is more appropriate for each set of shapes. Use any criterion of appropriateness you like or just go 'by feel.' Indicate your choice by drawing a line from the name to the box with the shapes. [Further down on the page:] State in your own words what makes your choice of names for the shapes seem appropriate to you."

---

[7] The expected frequencies for the chi-square computations were corrected for slight response biases in the overall frequencies of usage of the five trigrams.

A total of 49 new undergraduate *S*s filled out these sheets. Of these, 37 responded in the expected direction by choosing HAB for Shapes I, VII, and X, and ZEQ for Shapes III, IV, and IX. The probability of this occurrence is less than .01.[8] The most frequently stated rationale for the choices was in terms of perceived similarity between shapes and labels. Thus, 28 of the 37 *S*s assigned HAB to Shapes I, VII, and X because they perceived both the shapes and the label as "solid," "stable," "bulky," or "soft," while ZEQ and Shapes III, IV, and IX were perceived as "sharp," "light," "unstable," and "thin." That these dimensional categorizations have their measurable physical correlates in the stimuli is confirmed by the unit-area/squared-perimeter ratios ('density') of the shapes: .037, .026, and .024 for Shapes I, VII, and X, and .013, .018, and .013 for Shapes III, IV, and IX. *S*s responding in the opposite direction gave a greater variety of rationales, including remote associations with conventional words or objects and similarity between shapes and letters.

DISCUSSION

The results show that even with the most randomly selected materials, and in the absence of either meaningful referents or suggested dimensions, individuals consistently consider some sounds more appropriate than others as names for unfamiliar stimuli. Further, considerable agreement among subjects can be obtained for some label-referent combinations. This finding appears to contrast with previous interpretations (see nn. 4 and 5) which suggest that subjects need to be given a common referent for agreement to occur.

It might be argued that the consistent label-referent matching in this study could be an artifact of presenting each shape 10 times —that after his first choice of label for a particular shape, the individual could remember which label he used and respond with it again on later presentations of the same shape. However, their comments indicated that few subjects were aware that the shapes were being repeated. And even if they all knew this, recalling which of five labels was previously selected for which of 10 figures, all of them being presented in random order and at a rapid rate, would appear a remote possibility to account for the degree of consistency found in the data. Most importantly, while some consistency *within* each subject's responses might be attributed to memory on later trials, the agreement *among* subjects cannot be explained in this way, since agreement on later trials would have to be based on even greater agreement on the very first trials.

---

[8] Normal approximation of the binomial with large samples; see S. Siegel, *Nonparametric Statistics for the Behavioral Sciences*, 1956.

How can randomly constructed and selected materials produce reliable and consistent label-referent combinations? One possible explanation is in terms of an associative-probability theory. According to this general view, associates are covertly produced to both the CVCs and the random shapes. In those instances where the associations converge, a mediational link is provided and the CVC is chosen as the appropriate label. Subsequent presentations strengthen the initial association. From this hypothesis one might expect that the CVCs with the highest association values would be used most frequently and consistently as labels. However, of the five CVCs HAB and ZEQ have the lowest Glaze association values.[9] In the more recent list published by Archer,[10] HAB is 64% while ZEQ is 15%, this latter being the lowest association value for the five CVCs used in the study. The lack of relationship between association value of a trigram and its consistent use as a label agrees with previous findings that agreement among individuals assigning nonsense syllables to referents was not related to *m′* (meaningfulness as high association value).[11]

An alternative explanation of the findings is provided by the representational theory of reference, which holds that the process of naming involves an attention to and utilization of perceptual similarities between vehicle, *i.e.* vocal sound, and referent.[12] Wisseman has demonstrated that subjects can invent onomatopoeic words to represent novel sounds and that such inventions show both within- and between-subject consistency.[13] Onomatopoeia was not involved in the present study, but the representational theory maintains that dimensions of perceptual similarity can be found across perceptual modalities as well as within the same modalities. Consequently, various characteristics of vocal sounds (*e.g.* pattern, mode of articulation) can be used to represent attributes of objects.

---

[9] B. J. Underwood and R. W. Schulz, *Meaningfulness and Verbal Learning*, 1960.

[10] E. J. Archer, Re-evaluation of the meaningfulness of all possible CVC trigrams, *Psychol. Monogr.*, 74, 1960 (No. 497).

[11] Weiss (see n. 5); B. G. Rosenberg, P. Badia, and J. Langer, A representational assessment of meaningfulness, *Psychon. Sci.*, 1, 1964, 263–264.

[12] Werner and Kaplan (see n. 3).

[13] Quoted by Brown (see n. 1).

The associative-probability and the representational theories need not be mutually exclusive. Indeed, several subjects in the present study made use of converging associations to create the 'appropriate' linkage. Thus, one called Shape X a HAB because "HAB made me think of 'habla' [speak] and that shape looks like an open mouth." Nevertheless, the predominant rationale referred to shared attributes in the shapes and the labels. Thus, one subject stated that "HAB is a heavy-sounding word and these shapes seemed heavier," while another wrote that "the objects in the first box seem fuller and HAB seems to have more body."

A representational theory of reference does not deny that in conventionalized and tightly organized symbolic systems, as in a language, the relationship between the symbol and its referent is nearly arbitrary. Well-practiced users of such a system may well propose meanings for novel sounds or names for novel referents through association with already familiar sound-referent combinations. However, the theory does maintain that humans have a symbol-generating capacity which is based on perceived similarities between the symbol and the referent. This capacity is more evident in young children, who do not yet make use of a comprehensive and overlearned conventional language system. It can also become operative when older subjects accept a naming set in a situation which minimizes the opportunities for language-based associations. Such appears to have been the case in the present study. With both the 'naming' and the 'referent' materials random and unfamiliar, and in the absence of a conventional reference point or verbal-judgment dimension, most subjects were able to polarize both the shapes and the CVCs along some dimension of perceived similarities. It is clear from their comments that many of these subjects saw their chosen labels as *descriptive* of the referents rather than as something of which they were *reminded* by an associative chain.

KNOWLEDGE OF RESULTS AND ITEM ASSOCIABILITY IN PAIRED-ASSOCIATES LEARNING

Laraine Masters, *University of Illinois*

Abstract. The effects of percentage of knowledge of results (KR) and associability of items were investigated in the acquisition and extinction of paired associates. Two levels of each of these variables were factorially combined to produce four treatment groups of 24 undergraduates each. Both factors affected acquisition: 50% KR groups took longer to reach criterion than did 100% KR groups; low-associability groups took longer to reach criterion than did high-associability groups. Withdrawal of knowledge of results produced an increase in correct responses over trials for all groups. Natural language mediators were recorded and shown to facilitate accuracy during acquisition. Both acquisition and extinction were characterized by a dropping out of these mediators.

As a result of the attempt to subsume human learning under the already established paradigms of animal learning, knowledge of results has been assumed to operate as does empirical reinforcement, and much of the experimental work on the knowledge of results in paired-associates learning has attempted to deal with this independent variable by extrapolating from the techniques of animal learning. An examination of the evidence, however, indicates that the data obtained with animals are considerably different from those obtained with humans, for although acquisition proceeds more rapidly with both animals and humans under a schedule of 100% reinforcement than under one of partial reinforcement, resistance to extinction does not seem to be any greater for partially reinforced than for continuously reinforced humans.[1] This finding

---

Received for publication July 25, 1969. The paper is based on a thesis submitted to the University of Illinois in partial fulfillment of the requirements for the M.A. The author expresses her thanks to Dr. Jack A. Adams, whose able assistance was always available.

[1] A. E. Goss and E. J. Rabaioli, Response strength in a modified Thorndikian multiple-choice situation as a function of varying proportions of reinforcement, *J. exp. Psychol.*, 43, 1952, 106–114; A. E. Goss, C. H. Morgan, and S. J. Golin, Paired-associates learning as a function of percentage of occurrence of response members (reinforcement), *J. exp. Psychol.*, 57, 1959, 96–104.

contrasts with the partial reinforcement effect (PRE) shown to apply both to animal-conditioning and to simple human-conditioning situations. Furthermore, some investigators have found that withdrawal of knowledge of results during paired-associates learning leads not to extinction but rather to a continued increase of response strength over trials in which no knowledge of results is presented.[2] This effect has been called subjective or self-reinforcement.

Recent work has also demonstrated that the learning of paired associates is behaviorally more complex than it was once thought to be. As Battig has pointed out, humans use coding and grouping processes even in paired-associates tasks where extensive efforts have been made to eliminate cues for such coding.[3] Natural language mediators—early investigated and recently defined as implicit verbal chains imposed by the subject between stimulus and response[4]—have also been shown to be facilitative both in acquisition[5] and in retention[6] of paired associates.

Experiments which have manipulated percentage of knowledge of results as an independent variable have not been concerned with natural language mediation. This experiment attempted to integrate the two areas by investigating the percentage of knowledge of results and the associability of items in the acquisition and extinction of paired associates. The manipulation of item associability was designed to affect the formation of natural language mediators, with high associability leading to greater mediator formation than low associability.

---

[2] J. Richardson, The relationship of stimulus similarity and number of responses, *J. exp. Psychol.*, 56, 1958, 478–484; P. D. Eimas and D. Zeaman, Response speed changes in an Estes paired-associates "miniature" experiment, *J. verb. Learn. verb. Behav.*, 1, 1963, 384–388; J. Richardson and M. S. Gropper, Learning during recall trials, *Psychol. Rep.*, 1964, 551–560.

[3] W. F. Battig, Evidence for coding processes in "rote" paired-associate learning, *J. verb. Learn. verb. Behav.*, 5, 1966, 177–181.

[4] H. B. Reed, Associative aids: I, Their relation to learning, retention, and other associations, *Psychol. Rev.*, 25, 1918, 128–155; J. A. Adams, *Human Memory*, 1967.

[5] Reed (see n. 4); H. O. Kiess and W. E. Montague, Natural language mediation in paired-associate learning, *Psychon. Sci.*, 3, 1965, 549–550; A. R. Jensen and W. D. Rohwer, Verbal mediation in paired-associate and serial learning, *J. verb. Learn. verb. Behav.*, 1, 1963, 346–352.

[6] W. E. Montague, J. A. Adams, and H. O. Kiess, Forgetting and natural language mediation, *J. exp. Psychol.*, 72, 1966, 829–833; F. J. Boersma, R. C. Conklin, and J. E. Carlson, Effects of reporting associative strategies on the retention of paired associates, *Psychon. Sci.*, 5, 1966, 463–464.

METHOD

Design. Two levels of of each of two factors—percentage of knowledge of results (KR) and associability of pairs in the lists to be learned by the anticipation method—were factorially combined to yield four treatment groups: Group L 50 received 50% KR, low associability; Group L 100 received 100% KR, low associability; Group H 50 received 50% KR, high associability; Group H 100 received 100% KR, high associability. Each list contained 10 paired associates, CVC pairs selected from a larger pool of items used by Montague and Kiess, who scaled the pairs for the number of natural language mediators (NLMs) they elicited and determined an associability value from the proportions of *S*s giving an NLM to the pair.[7] In the present study, the items on the high-associability list had a mean scale value of 91.3%; it was 20.2% for the items on the low-associability list.

The experiment was divided into an acquisition and an extinction phase. Extinction procedures were the same as those used in acquisition except that all KR was withdrawn for all groups. In both acquisition and extinction, *E* recorded the accuracy of each response as well as whether or not an NLM was used to recall that response. Four orders of list presentation were used. Half of the *S*s in each group started acquisition with Order 1 and half started with Order 3. The beginning order of extinction for each *S* was the one that had not been used for the start of acquisition. The determination of order of items was random with the stipulations that no item appear in the same position more than once and that no item precede or follow any other item more than once.

The KR contingencies for the two 50% KR groups were the same. They were randomly determined with the qualifications that the 50% level of KR be kept within each order and for items across orders and that no more than three items in succession receive or not receive KR.

Procedure. The *S*s were instructed on the use of the anticipation method, modified to permit the reporting of NLMs as well as responses. The *S*s were told to report any mediator they had used immediately after they had given the response and before KR was presented. If no mediator had been used, *S*s were told to say 'none.' Acquisition proceeded to a criterion of 6/10 correct for the L groups and 5/10 correct for the H groups. (A pilot study had indicated that *S*s getting the items from the high list learned so quickly that if they had 6/10 correct on the last acquisition trial, the learning on that final trial often brought them to a final acquisition level of 8/10 or 9/10 correct. Thus, the criterion for the H groups was lowered on the assumption that they would learn at least one more pair on the final acquisition trial, thus making them equivalent to the L groups in their final performance level.) If *S* did not reach criterion in 18 trials, the extinction phase was initiated nonetheless. After acquisition was complete, *S* was given a brief rest and then 18 trials of extinction were run.

[7] W. E. Montague and H. O. Kiess, The associability of CVC pairs, *J. exp. Psychol.*, 78, 1968 (No. 2, Pt. 2), 1–23.

Apparatus. The stimuli were presented by a Kodak Carousel projector. Two successive slides made up a presentation on a trial. The first slide was the stimulus alone for anticipation, and the second slide had the S-R pair for a reinforced trial or a blank slide for a nonreinforced trial. The presentation rate was 4:4. The projected letters were approximately 3 in. high. The *S* sat 4½–5 ft. away from the wall on which the letters were projected.

Subjects. Ninety-six undergraduates (24 women, 72 men) in an introductory psychology course served as *S*s. In each of the four treatment groups, 6 women and 18 men served as *S*s. Within this restriction, *S*s were generally assigned to successive treatment groups in order of their appearance in the lab.

RESULTS

Acquisition. The number of trials to criterion was the number of learning trials it took *S* to get 6/10 correct (or 5/10 correct for the H groups) on one trial. If *S* did not reach criterion after 18 trials, his score was determined by adding one trial for every item fewer than criterion; *e.g.* an *S* in an L group who had 4/10 instead of 6/10 correct on Trial 18 would have a score of 20. Ten *S*s in the L 50 group had their scores corrected in this manner: four *S*s had 5/10 correct; three *S*s had 4/10 correct; two *S*s had 3/10 correct; and one *S* had only 2/10 correct. In the L100 group, two *S*s did not reach criterion: both had only 3/10 correct and, thus, scores of 21. No corrections were needed for any *S* in either of the H groups. An extra trial was also added to the scores of those *S*s in the H groups who reached only 5/10 correct, in order to more closely equate those groups with the L groups who went to a criterion of 6/10 correct. Table I presents the means of number of trials to criterion for the four groups. The H groups took fewer trials to criterion than did the L groups. The 100% KR groups took fewer trials to criterion than did the 50% KR

TABLE I

MEANS FOR THREE PERFORMANCE MEASURES IN ACQUISITION

Group	Trials to criterion	Mean percentage correct	Percentage items mediated
L 50	15.4	24.7	27.9
L 100	12.7	27.2	24.6
H 50	3.8	46.6	39.6
H 100	2.6	57.7	37.1

groups. A 2 × 2 fixed-factors analysis of variance indicated significance for associability value [F (1, 92) = 216.44, $p < .001$] and for percentage KR [F (1, 92) = 6.84, $.025 > p > .01$], but not for their interaction. A calculation of ω^2 indicated that the associability value of items was by far the strongest factor for this variable, accounting for 68% of the variance.

The mean percentage of items correct (computed by dividing the total number of correct responses over all trials by the number of opportunities to be correct over all trials; converting that quotient into percentages; and averaging the percentages within groups) was, it should be pointed out, uninfluenced by the addition of trials to the scores of *S*s who did not reach criterion in 18 trials: the number of opportunities to be correct could never exceed 180 (18 trials × 10 items per trial) no matter how many trials were added to the scores. It is interesting, therefore, that the results for mean percentage correct (Table I) indicate much the same thing as did the trials to criterion. Percentage KR had some effect, since mean percentage correct was lower for the 50% KR groups in both the H and L conditions. Once again, however, associability value seemed to be a stronger factor. An analysis of variance substantiated these observations: associability value [F (1, 92) = 178.99, $p < .001$] and percentage KR [F (1, 92) = 12.87, $p < .001$] were significant, as was their interaction [F (1, 92) = 4.07, .05 $p >$ .025].

A measure of whether each item was mediated or rote was obtained for each item for each trial. An item was defined as mediated if it had a mediator on any trial, regardless of the number of times the mediator reappeared thereafter. Table I presents the mean percentage of mediated items for each group. The H conditions had a higher percentage of mediated items than did the L conditions, and for both these conditions, *S*s receiving 50% KR had a slight tendency to mediate more than *S*s receiving 100% KR. Associability value was found to have a significant effect [F (1, 92) = 5.65 $.05 > p > .025$], but neither percentage KR nor its interaction with associability value was significant as tested by a 2 × 2 fixed-factors analysis of variance.

Additional information on the use and function of mediators in the acquisition phase was obtained by classifying responses as to whether they were first correct before a mediator was reported, first correct at the same trial that a mediator was first reported, or first correct after (or never correct even after) a mediator was reported (Table II). For all four groups, the large majority of responses was correct for the first time when a mediator was first reported. Also, for all groups, an item that had a mediator had a much higher probability of being correct on criterion than one that did not have a mediator. For the *S*s who mediated, the percentage of mediated items incorrect at criterion was 10.2%. The percentage of rote items incorrect at criterion was 54.4% over all *S*s.

TABLE II

PERCENTAGES OF RESPONSES FIRST GIVEN CORRECTLY BEFORE, DURING, OR AFTER A MEDIATOR WAS FIRST REPORTED FOR AN ITEM

Group	Before	During	After
L 50	24.2	68.2	7.6
L 100	28.6	64.3	7.7
H 50	19.7	75.5	5.3
H 100	10.3	80.4	9.2

Finally, the use of mediators did not remain constant in acquisition but, rather, decreased as acquisition progressed. The percentage of items mediated in acquisition was 26.2% and 33.3% for the L and H groups respectively, while the percentage of mediators used in the last acquisition trial was 14.8% and 26.2% for the L and H groups respectively.

Extinction. The 18 extinction trials were combined into six blocks with three trials in each block for the analyses of both mean number of correct items and mean number of mediators. Table III presents the means for the six blocks for each treatment group. In both cases, the effect of blocks was significant, with the mean number of correct items increasing over blocks [F (5, 460) = 53.00, $p <$ .001] and the mean number of mediators decreasing over blocks [F (5, 460) = 21.02, $p <$ .001]. In addition, the H groups had significantly more correct responses than the L groups [F (5, 460) = 18.31, $p <$ .001]. The interaction between associability value and blocks [F (5, 460) = 3.15, $p <$ .01] was significant, showing that the L groups improved more over blocks than did the H groups. These last two tests, however, are confounded by the differences in the associative strength for the high- and low-associability groups as measured by number of correct responses at the end of ac-

TABLE III

Effect of Extinction on Numbers of Correct Items and Number of Mediators

	Mean number correct						Mean number mediators					
	Blocks						Blocks					
Groups	1	2	3	4	5	6	1	2	3	4	5	6
L 50	5.4	5.7	5.9	6.2	6.3	6.4	1.7	1.4	1.3	1.2	1.2	1.1
L 100	5.7	6.4	6.6	6.8	6.8	6.9	1.3	.9	.7	.5	.4	.3
H 50	7.0	7.4	7.5	7.6	7.6	7.6	2.8	1.9	1.5	.90	.94	.88
H 100	6.8	7.5	7.7	7.8	7.8	7.9	2.3	1.9	1.6	1.3	1.2	.99

quisition. The mean number correct for the H groups on the first extinction trial was 7.02; it was only 5.58 for the L groups. The difference of 1.44 was significant at $p < .001$ ($t = 17.14$, $df = 94$). Neither percentage KR [F (1, 92) = 1.77] nor the interaction of percentage KR with any of the other factors was significant. For number of mediators, neither associability value nor percentage KR was significant, but both blocks (F reported above) and the interaction of associability value and blocks [F (5, 460) = 3.06, $p < .01$] were significant. This interaction is uninterpretable for the same reason as was the similar interaction for number correct.

To better characterize the items that showed increments in response strength over extinction trials, *i.e.* that were incorrect on the first extinction trial but correct at some later time during extinction, they were classified as correct or incorrect in acquisition and as mediated or rote. An item was considered correct if it had been correct on at least one acquisition trial, and it was considered mediated if *S* reported a mediator on at least one acquisition trial. That trial did not have to be the criterion trial in either case. This analysis showed that a mediated item was more likely to show an increment in extinction if it had been correct at least once in acquisition than if it had never been correct. Of the mediated items that showed an increment, 79.2% were correct in acquisition and only 20.8% were incorrect. The same trend was not observed for rote items. Of the rote items that showed an increment, 44.6% were correct and 55.4% were incorrect. Within the rote items, however, different patterns were observed for the low- and high-associability groups: an approximately equal number of correct and incorrect items showed increments in extinction in the low-associability groups, while twice as many incorrect as correct items showed increments in extinction in the high-associability groups. These data must be interpreted in light of the fact (mentioned in the section on acquisition, above) that there were fewer mediated items incorrect at criterion (10.2%) than rote items incorrect at criterion (54.4%). Thus, there were very few mediated items that had never been correct; and it is not surprising that of the mediated items showing an increment in extinction, more were correct than incorrect.

Another general observation must be interpreted with caution. Of the 116 items showing an increment in extinction, 20.7% were mediated and 79.3% were rote. However, since only 32.3% of all items were mediated, the 20.7% is not as much below expectation as might have been thought.

DISCUSSION

Acquisition. As expected, both associability and percentage KR significantly affected acquisition. And, as would be expected, high-associability groups had more natural language mediators than did low-associability groups; the high-associability groups also learned more quickly. Such evidence supports much of the published research indicating the facilitative effect of natural language mediators in acquisition.[8] Similarly, the finding that 100% KR resulted in fewer trials to criterion than did 50% KR corroborates the majority of studies which have manipulated percentage of knowledge of results as an independent variable.[9]

There is no evidence that percentage of knowledge of results affected formation of natural language mediators. While the percentage of mediators was significantly higher for the H groups than for the L groups, no differences between 50% and 100% KR were found to be significant. The degree of formation of the mediators remained fairly constant in spite of changes in percentage of knowledge of results. When mediators were formed, however, they appeared to help maintain accuracy. Two results of this

[8] L. L. Clark, T. G. Lansford, and K. M. Dallenbach, Repetition and associative learning, *J. exp. Psychol.*, 73, 1960, 22–40; Jensen and Rohwer (see n. 6 above).

[9] Goss and Rabaioli (see n. 1 above); R. W. Schulz and W. N. Runquist, Learning and retention of paired adjectives as a function of percentage and occurrence of response members, *J. exp. Psychol.*, 59, 1960, 409–413.

experiment support such a conclusion. (1) Of the responses 72.1% were first correct at the same time that mediators were first reported, suggesting that response strength is, in some way, dependent on mediation. (2) An item that had a mediator was more likely to be correct on criterion than one that did not have a mediator: for all groups, only 10.2% of mediated items were incorrect on the criterion trial, whereas 54.4% of the rote items were incorrect at criterion.

Extinction. The data for the extinction phase of the experiment manifested not only the lack of extinction but also the presence of self-reinforcement, an increase in correct responses over blocks of trials. These results agree with those of such studies as Eimas and Zeaman's, Richardson and Gropper's, and Greenbloom and Kimble's.[10] No differences in performance over blocks, as measured by mean number of correct responses, were found for the percentage-knowledge-of-results variable. The significant interaction between associability and blocks of trials was probably an artifact of the L groups' starting extinction at a lower level and thus having more room for improvement.

The significance of blocks for number of mediators indicates extinction of mediators over trials. Actually, the dropping out of mediators started in acquisition, since the percentage of items mediated in acquisition was 26.2% and 38.3% for the L and H groups respectively, while the percentage of mediators used in the last acquisition trial was 14.8% and 26.2% for the L and H groups respectively. Thus, the mediators were used for a certain time; then they dropped out and the item was recalled in a rote fashion. These data support those of Adams and McIntyre, who also found that mediators drop out during acquisition trials.[11] The failure of the interaction of blocks and percentage knowledge of results to reach significance for number of mediators in extinction indicates that this dropping out is independent of manipulations of the percentage of knowledge of results during acquisition.

The fact that nearly half of the items showing self-reinforcement effects had never been correct in acquisition is one for which any tenable theory of verbal learning must account. Adams (see n. 4) relates self-reinforcement to a closed-loop theory in which there exists the possibility for internal reinforcing events after external reinforcement has placed the response in the repertoire. Instead of comparing the stimuli from the feedback loops of the response just made with the external knowledge of results, the subject compares them with internal reference levels. Such an explanation may be adequate for the case when the response has been given correctly at some time and, therefore, is placed in the subject's repertoire. However, it does not explain the case when the correct response has never been made during acquisition, unless it is assumed that the subject makes the response covertly but does not say it aloud because of a high response-criterion bias.

---

[10] Eimas and Zeaman (n. 2); Richardson and Gropper (n. 2); R. Greenbloom and G. A. Kimble, Extinction of R-S association and performance on recall trials without informative feedback, *J. verb. Learn. verb. Behav.*, 4, 1965, 341–347.

[11] J. A. Adams and J. S. McIntyre, Natural language mediation and all-or-none learning, *Canad. J. Psychol.*, 21, 1967, 436–449.

ELECTROMYOGRAPHIC ACTIVITY DURING OBSERVATIONAL LEARNING

SEYMOUR M. BERGER, DEBORAH SMITH IRWIN, and GABRIEL P. FROMMER,
Indiana University

Abstract. A previous study used learning of hand signals to show that observers practice the model's responses during the exposure period. Since only overt practice was measured, it was felt that more sensitive techniques were needed to measure possible covert practice. In this study, electromyogram (EMG) recordings were made while observers learned hand signals and word-number pairs, either in the presence of the experimenter or alone. Observers showed greater EMG activity while learning hand signals than while learning word-number pairs regardless of whether the experimenter was present or not. However, more overt practice of the hand signals (hence larger EMGs) occurred when the experimenter was absent. The results are interpreted as additional evidence for the elicitive function of the model's behavior during observational learning.

Typically, observational learning studies involve two stages: a stage in which an observer is exposed to the behavior of a model, and then a stage in which the observer is required to behave in order to determine how much the model's behavior has influenced the observer's behavior. Relatively little research has attempted to analyze the observer's reactions to the model's behavior during the exposure stage—even though identification of these reactions could, presumably, facilitate more precise analysis of the processes involved in observational learning. Still, some efforts in this direction have been made,[1] and it has, for example, been found that observers

---

Received for publication June 10, 1969. The study is based in part on Deborah Smith Irwin's undergraduate honors thesis. It was supported by National Science Foundation Grant GS-1702 to the first author. Send reprint requests to the first author at the Department of Psychology, University of Massachusetts, Amherst, Mass. 01002.

[1] A. Bandura and T. L. Rosenthal, Vicarious classical conditioning as a function of arousal level, *J. pers. soc. Psychol.*, 3, 1966, 54–62; S. M. Berger, Conditioning through vicarious instigation, *Psychol. Rev.*, 69, 1962, 450–466; V. Di Lollo and S. M. Berger, The effects of apparent pain in others on observer reaction time, *J. pers. soc. Psychol.*, 2, 1965, 573–575; E. W. J. Faison, N. Rose, and J. E. Podell, A technique for measuring observable audience reactions to training films, Air Force Personnel and Training Research Center, Training Aids Research Laboratory, unpublished laboratory note (mimeo.), TARL-LN-55-45, 1955.

who watched a model learn hand signals (from the manual alphabet for the deaf) tended voluntarily to perform these hand signals as they watched and that the performance of hand signals during the exposure period was moderately correlated with a measure of learning (.47).[2]

The performance of the hand signals in that study was, however, scored by a relatively strict criterion, which fragmentary performance or unidentifiable hand movements did not meet; it was apparent that more hand activity was occurring than was being scored. The purpose of the present study, then, was to determine the extent to which exposure to hand signals increases the amount of motor activity in the observer's hand and arm. Records of overt activity were obtained as in the previous study, but in addition, electromyogram (EMG) recordings were obtained to measure any less overtly visible motor activity which might occur. The rationale for using EMG recordings was based on the evidence that thinking is related to local muscular changes.[3] So, for example, thinking about moving one's own hand is related to changes in the electrical potentials of muscles in that region of the body. These changes could be produced by self-instruction. It seemed reasonable to hypothesize that watching another person's hand movements could have an effect similar to the one produced by self-instruction, *i.e.* that local muscular changes would be produced in an observer as a consequence of his watching a model's performance. This seemed especially reasonable given the earlier study's evidence of overt observer performance.

The study was conducted in two phases. In the first phase, the experimenter was not present in the observer's section of the room; in the second phase, she was. The experimenter's presence was introduced as a factor because pilot work for the earlier study suggested that observers tended not to perform hand signals in the experimenter's presence but did practice when apparently alone. (They seemed to feel that performing hand signals overtly is akin to counting on one's fingers—an act thought to be embarrassing for an adult.) The effect of the experimenter's presence, therefore, was expected to reduce the amount of overt practice of the hand signals. Finally, since the EMG activity observed might not be specific to the hand signals but the result of a general activation during learning, all observers were given another learning task, one which was not associated with the hand. This comparison task involved word-number pairs.

---

[2] S. M. Berger, Observer practice and learning during exposure to a model, *J. pers. soc. Psychol.*, 3, 1966, 696–701.

[3] E. Jacobson, The electrophysiology of mental activities, this JOURNAL, 44, 1932, 677–694; M. A. Max, An experimental study of the motor theory of consciousness: III, Action-current responses in deaf-mutes during sleep, sensory stimulation, and dreams, *J. comp. Psychol.*, 19, 1935, 469–486; M. A. Max, Experimental study of the motor theory of consciousness: IV, Action-current responses of the deaf during thinking, *J. comp. Psychol.*, 24, 1937, 301–344.

METHOD

Subjects. The *S*s who served as 'observers' were 31 students in elementary psychology courses at Indiana University. They participated in order to fulfill a course requirement.[4]

Apparatus. EMGs were picked up with Grass silver-disc electrodes coated with Sanborn electrode jelly and secured to *S*'s right (preferred) arm with adhesive stretch tape. The muscle potentials were amplified and recorded with a Grass Model 7 polygraph. In one channel, the signal was integrated with a Grass Model 7P3 preamplifier-integrator, and only the integrated signal was available for the write-out. In the other two channels, the EMGs were amplified with a Grass Model 7P5 preamplifier set to a low-frequency cutoff of 10 Hz. The preamplified signal was fed to a driver amplifier and the raw record was displayed. In addition, the raw signal was picked up from J6 of the driver amplifier, half-wave rectified, and fed to the input of another driver amplifier, the high-frequency cutoff of which was set at 3 Hz. This configuration functions on the same 'leaky-condenser' principle described by Arduini and Pinneo, and others, which provides, in effect, a sliding average of the input signal.[5] These summated signals were displayed in parallel to the raw records, and the measurements reported here were made on these summated signals.

Pictorial representatives of eight dissimilar hand signals from the manual alphabet for the deaf were mounted on 3¾ × 5¾ in. white paper with the corresponding letter of the alphabet typed beneath each signal. These hand signals represented the behavior of the 'model.'[6] Eight common five-letter English words paired with numbers from 1 through 8 were typed on another 3¾ × 5¾ in. white paper. The room was divided by a wooden screen which separated *S* from the polygraph and in one condition, from *E*.

Procedure. Each *S* was run individually. He was seated in a large-armed chair and told that harmless electrodes were going to be attached to his arm and that no shock was to be administered. Three sets of electrodes were then applied as follows: (1) wrist position—over portions of the flexor carpi radialis, palmaris longus, and flexor digitorum superficialis muscles, one electrode placed slightly above the wrist, and the second 1¼ in. above the first; (2) forearm flexor position—over the flexor carpi radialis and flexor digitorum sublimis muscles, one electrode placed one-third of the distance from the medial humeral epicondyle to the styloid process of the radius, the other electrode 1¾ in. in the distal direction along the same line; (3) forearm extensor position—over the superficial extensor muscles of the forearm, one electrode placed one-third of the distance from the lateral humeral epicondyle to the styloid process of the ulna, the other electrode 1¾ in. in the distal direction along the same line.

---

[4] More students participated in part or all of the experiment, but their data had to be eliminated because of equipment failures.

[5] A. Arduini and L. R. Pinneo, A method for the quantification of tonic activity in the nervous system, *Arch. ital. Biol.*, 100, 1962, 415–424.

[6] Static pictorial representations of the model's behavior were used rather than a 'live performance' in order to eliminate the possibility that the observers' EMGs were instigated by the model's irrelevant arm movements rather than his performance of the hand signal per se.

When this attachment was complete, the following instructions were read to the *S*s who were to learn the hand signals first; the instructions were the same for the *S*s who learned the word-number pairs first, except for a reversal of the order in which the learning tasks were described.

> "In this experiment you will be learning two different tasks. One task will be to learn the eight hand signals from the manual alphabet for the deaf that correspond to the correct eight letters. Are you familiar with any of these signals? [Any *S* who could reproduce any of the signals was excused.] An example of such a signal is K. [*E* demonstrates the signal for K.] The other task will be to learn eight word-number pairs; *i.e.* for each word there will be associated a number from 1 through 8. You are to learn the number that corresponds to each word. An example of this task would be TABLE-9.
>
> "Do you have any questions? For the first task you will be tested in the following way. I will give you a letter such as K and you will demonstrate the corresponding hand signal. For the second task I will give you a word, *e.g.* TABLE, and you will tell me the correct corresponding number, *i.e.* 9. Do you understand how you will be tested? There will also be three relaxation periods in this experiment: one before the first learning task, one after the first task, and one after the second task. During these periods please close your eyes and remain as relaxed as possible. When I say you may now start learning the first task, open your eyes. The task materials will then be in your lap. Learn the tasks as quickly as possible. When you think you have learned all eight items, tell me and I will then test you on that task. Do you have any questions about what you are to do? If not, then close your eyes and the experiment will begin."

At this time *E* turned on the EMG console, adjusted baselines when necessary, calibrated the incoming signal for each channel, and began recording the EMG. Either after a period of approximately 45 sec. or when the *S* had sufficiently relaxed (to establish baseline level of EMG activity), *E* put the materials for one of the two learning tasks on each *S*'s lap (half the *S*s learned the hand signals first while the other half learned the word-number pairs first), and proceeded as the instructions had promised.

E indicated each rest and learning period on the polygraph paper by operating an event marker, and also recorded the number of correct responses *S* made in the test session. At the close of the experiment *S* was excused. If asked, *E* then briefly explained the experiment.

There were two conditions involving *E*'s presence: in Condition I *E* was in front of *S*, and in Condition II *E* was behind a screen adjacent to *S*. *E* could see *S*'s right hand from this position without being seen by *S*. The statistical analyses are based on data from both of these phases of the experiment.

Muscle-action potentials for each *S* were measured in arbitrary units from the zero-input baseline at 5-sec. intervals for the three rest periods and the two learning periods. Analysis was performed only on data obtained during the learning periods. Although muscle-action potentials were recorded at three positions, the data from one position were unusable due to technical difficulties. Consequently, only data from these two positions were analyzed: Channel I, wrist position, and Channel II, forearm flexor position.

RESULTS

Data from Channel I (wrist position) were obtained for Conditions I and II, but only the data from Channel II (forearm flexor) under Condition II were available for analysis because of technical problems with this recording channel during Condition I. The main data analysis, therefore, is based upon Channel I recordings.

The mean muscle-action potentials recorded on Channel I are reported in Table I. The analysis of variance performed on these data from 10 *S*s run under Condition I and 10 *S*s under Condition II yielded two significant results. First, there was a significant main effect due to the type of learning task; *i.e.* there was a significantly greater amount of muscle activity during the hand-signal learning than during the word-number learning ($p < .01$). Second, there was a significant interaction between the conditions associated with *E*'s presence and the type of learning task ($p < .01$). No main effect was found for conditions involving *E*'s presence as such. The analysis of variance is summarized in Table II.

The significant interaction in the analysis of variance shows that

TABLE I

MEAN ELECTROMYOGRAPHIC ACTIVITY SCORES IN ARBITRARY UNITS

	Learning task		
Condition	Hand-signals	Word-number pairs	Total
I	37.32	11.84	49.16
II	21.82	15.92	37.68
Total	59.14	27.76	

TABLE II

SUMMARY OF ANALYSIS OF VARIANCE OF CHANNEL I SCORES

Source	*df*	*MS*	*F*
Conditions	1	326.21	1.65
Error (a)	18	197.75	
Type of task	1	2462.23	35.52*
Conditions × Task	1	958.54	13.83*
Error (b)	18	69.32	

* $p < .01$.

the magnitude of the effect of the learning task depended on the presence of *E*. Separate *t* tests comparing different means showed that the mean activity associated with hand-signal learning was greater under Condition I, when *E* was absent, than under Condition II, when *E* was present ($t = 2.78$, $df = 18$, $p < .02$). On the other hand, the same comparison for word-number learning showed that the means for the two conditions were not significantly different ($t = 0.86$, $df = 18$). Within both conditions, however, more activity was associated with learning hand signals than with learning word-number pairs (for Condition I, $t = 5.96$, $df = 9$, $p < .01$; for Condition II, $t = 3.01$, $df = 9$, $p < .02$).

The *S*s were also scored by *E* during hand-signal learning either for performing or not performing overt recognizable signals (see Table III). Those who did not perform such recognizable signals included the *S*s from whom fragmentary movement or no overt movement was recorded. The Fisher exact-probability test showed that a significantly greater amount of overt practice occurred under Condition I than under Condition II during hand-signal learning ($p = .01$).

Turning to the results from Channel II, a *t* test performed on data for 19 *S*s run under Condition II was also found to be significant (mean during hand-signal learning, 15.14; mean dur-

TABLE III

NUMBER OF OBSERVERS PERFORMING OVERTLY

Condition	Performers	Nonperformers
I	9	1
II	2	8

ing word-number learning, 8.63; $t = 3.11$, $df = 18$, $p < .01$), providing evidence from a different recording position of a significantly greater amount of activity which could not be detected by direct observation by *E* during hand-signal learning than during the control task.[7]

Finally, no significant differences were found in learning per se. It took *S*s just about as long (in terms of 5-sec. units) to learn the words as to learn the signals, and they learned as many words as they did signals. These results were the same within both conditions and there was no difference between conditions.

DISCUSSION

The principle findings of this investigation can be summarized as follows. (1) There was a significantly greater amount of muscle activity recorded during hand-signal learning than during word-number learning, regardless of whether or not the experimenter was present. (2) The amount of muscle activity during hand-signal learning under the condition where the experimenter was absent was significantly greater than under the condition where she was present. Such activity was not significantly different under the two conditions for word-number learning; thus, the interaction between the types of learning tasks and the conditions involving the experimenter's presence was significant. (3) A significantly greater proportion of observers performed overt, recognizable signals when the experimenter was absent than when she was present.

The first finding is consistent with related studies.[8] More specifically, it suggests that observers did not merely passively watch the model (here represented by pictures of the hand signals) during the learning period. Even though such active participation was not always at the overt behavioral level, increased activity presumably related to the hand-signal learning was revealed for most observers regardless of whether the experimenter was present or not. This ties in with the recent neurophysiological evidence suggesting an important role for anticipatory neuromuscular adjustments in the acquisition of simple conditioned responses. Integrated overt contraction of skeletal muscles normally requires the action of a complex sensory-feedback system. Phasic or tonic discharge of the gamma efferent motor neurons produces contractions of muscle-spindle stretch receptors but does not result in measurable overt movement. Rather, it increases the afferent bombardment from the stretch receptors onto the alpha motor neurons, the discharge of which initiates muscle contraction and movement.[9] Buchwald and co-workers have investigated the role of this feedback system in conditioned leg flections of the cat.[10] They showed that the gamma efferent discharge becomes conditioned well before conditioned alpha motor-neuron responses are elicited and overt movements appear. Further, overt CRs fail to develop on test trials given before blocks of extended training trials during which Flaxedil is used to block entirely the contraction of the muscle spindles induced by gamma efferent activity. It might also be suggested here that had electrical recordings been taken from regions such as the jaw and throat, they could have been expected to show increases in muscle-action potentials during the word-number learning, due to the vocal and subvocal responses previously found during a rote-learning task in a related study.[11]

---

[7] The 19 *S*s in this comparison for Channel II included 8 *S*s from the previous analysis of Channel I data. The additional 11 *S*s in the Channel II analysis brought the total number of *S*s in this study to 31.

[8] Berger (n. 2 above); Jacobson, and Max (first two references in n. 3 above).

The second and third findings, that the presence of the experimenter had a significant inhibitory effect on overt practice of the hand signals, had been previously observed but not tested (see n. 2). Yet, even when the experimenter was present, the electromyogram records showed greater activity during hand-signal learning than during word-number learning. Such evidence indicates that less obvious forms of signal practice probably occur during observational learning. Whether overt and recognizable practice with the hand signals was more advantageous for learning than fragmentary or no overt practice could not be tested in the present experiment because of the unlimited time allotted for learning. No differences in learning were found, but design of the experiment did not allow for the use of more sensitive measures of learning.

---

[9] P. B. C. Mathews, Muscle spindles and their motor control, *Physiol. Rev.*, 44, 1964, 219–288.

[10] J. S. Buchwald, D. Beatty, and E. Eldred, Conditioned responses of gamma and alpha motoneurons in the cat trained to conditioned avoidance, *Exp. Neurol.*, 4, 1961, 91–105; J. S. Buchwald, D. Beatty, and E. Eldred, Distribution and specificity of gamma motoneuron conditioned responses, *Exp. Neurol.*, 6, 1962, 524–537; J. S. Buchwald and E. Eldred, Conditioned responses in the gamma efferent system, *J. nerv. ment. Dis.*, 132, 1961, 146–152; J. S. Buchwald, M. Standish, E. Eldred, and E. S. Halas, Contribution of muscle spindle circuits to learning as suggested by training under Flaxedil, *EEG clin. Neurophysiol.*, 116, 1964, 582–594.

[11] R. N. Berry and R. C. Davis, Muscle responses and their relation to rote learning, *J. exp. Psychol.*, 55, 1958, 188–194.

It is not possible, of course, to determine the exact nature of the covert muscle activity reported in this study. For instance, there is no way of determining whether the observation of different hand signals produced different patterns of changes in muscle activity. In other words, it is not clear that observers were practicing specific hand signals. It is clear, however, that a considerable amount of local muscle activity occurred. It seems reasonable, at this point, to conclude that exposure to the hand signals did produce local muscular changes which *at least prepared* the observer to perform the hand signals on his own. These muscular changes, therefore, are viewed as evidence that exposure to the model's performance *readies* the observer to perform the model's response and in this way facilitates observer practice and, possibly, learning. These results are interpreted as further evidence for the elicitive function of a model's behavior in observational learning.[12]

Although the results of this experiment clearly show that overt practice and covert muscle activity occurred during exposure to learning materials representing the model's behavior, we still need studies that apply such sensitive techniques to other forms of a model's behavior. It also would be useful to study the relationship between overt and covert forms of practice, and to determine the significance of such practice for observational learning.

---

[12] A. Bandura, Vicarious processes: A case of no-trial learning, in L. Berkowitz (ed.), *Advances in Experimental Social Psychology,* 2, 1965, 1–55; S. M. Berger, Vicarious aspects of matched-dependent behavior, in E. C. Simmel, R. A. Hoppe, and G. A. Milton (eds.), *Social Facilitation and Imitative Behavior,* 1968, 169–185.

MAGNITUDE ESTIMATION OF AVERAGE LENGTH: A FOLLOW-UP

Arthur L. Miller, *Marquette University*
Virginia M. Pedersen, *Southern Illinois University*
and Richard W. Sheldon, *George Washington University*

Abstract. Miller and Sheldon have fitted power functions to magnitude estimations of average length of sets of horizontal lines and have found exponents of 1.00–1.14, depending on the standard. The present study obtained magnitude estimations of average length under quite different observation conditions, with different constraints on observers' judgments, and with different observer populations. These variations in experimental conditions and populations produced only small, and not entirely consistent, differences in slopes (exponents), and the variation in constraints had no consistent influence. As in the Miller and Sheldon study, a 12-in. standard yielded relatively steep slopes; a 60-in. standard, flat slopes. Differences between three groups with continuously visible standards were not, however, statistically reliable, as they were in the Miller and Sheldon study. Exponents of these three groups plus two low-constraint groups were found to be .92–1.03; most exponents clustered around 1.00–1.03.

In several recent psychophysical experiments, the observers have been given unusual tasks. In contrast to usual procedures, where the observers judge some attribute of unitary objects or events (*e.g.* the heaviness of a weight or the pitch of a tone), the observers in these experiments judged the *average* of a *set* of stimuli in terms of some attribute of the set's components. Anderson, for example, exposed sets of six weights and secured judgments of average heaviness on a 20-point scale.[1] Miller and Sheldon presented Army enlisted men with sets composed of six lines arranged in either parallel (horizontal) or oblique arrays and obtained magnitude

---

Received for publication April 24, 1969. The research was supported in part by a grant from the United States Public Health Service (MH-13397-01). The study was presented before the 1968 convention of the American Psychological Association.

[1] N. H. Anderson, Application of a weighted average model to a psychophysical averaging task, *Psychon. Sci.*, 8, 1967, 227–228.

estimations of average length of parallel, and average inclination of oblique, arrays.[2] While Anderson was primarily interested in the effects on judged average of serial position of stimuli within sets, Miller and Sheldon attempted to determine magnitude scales of average length and average inclination.[3]

In the study by Miller and Sheldon, the average for the array was judged in proportion to a standard line present on every trial, and each observer was informed of the physical value of the standard. Lines of three lengths (12, 36, and 60 in.) and three inclinations above horizontal (20°, 45°, and 70°) served as standards, and a given observer received only one length and one inclination standard. The stimulus arrays consisted of six black lines projected on a white screen and were exposed for 10 sec. each. The screen was viewed through a monocular occluder, and the visible sector was clearly bounded by the occluder aperture. The results showed the relationship between subjective magnitude of average inclination and stimulus magnitude to be linear, and this relationship for average length to be described by power functions with exponents (slopes in log-log coordinates) of 1.14, 1.00, and 1.03 for the standards of 12, 36, and 60 in. respectively.

Frequently in magnitude-estimation experiments, a standard stimulus is exposed or available to each observer on every trial, but often the first stimulus in the series is considered the standard and is not exposed on subsequent trials. The modulus (number assigned the standard) is often chosen by the observer or may be designated by the experimenter. Presenting the standard only once and allowing the observer to choose the modulus reduces the constraints on his judgments and may enhance the effectiveness of magnitude scaling (see n. 3). Since Miller and Sheldon exposed the standard on every trial and prescribed the modulus, one objective of the present investigation was to secure magnitude estimations with reduced constraints on the observers. A second objective was to ascertain whether relationships observed under the conditions of the Miller and Sheldon study would hold over different experimental conditions and populations. Our attention was restricted in this study to scaling average length.

---

[2] A. L. Miller and R. W. Sheldon, Magnitude estimation of average length and average inclination, *J. exp. Psychol.*, 81, 1969, 16–21.

[3] S. S. Stevens, The direct estimation of sensory magnitudes: Loudness, this JOURNAL, 69, 1956, 1–25.

METHOD

Observers. The *O*s were 80 paid volunteers (men and women) at the University of South Dakota. They demonstrated understanding of arithmetic means on a simple mathematical test and were unpracticed in psychophysical observation.

Apparatus and stimuli. The experiment was conducted in a darkened room containing a black projection screen (10½ × 12 ft.); a projector stand, located 26 ft. from the screen, housing a Kodak Carousel 35-mm. projector (Model AV-900); a desk (for *E*) beside the projector stand; and an armless chair (for *O*) 21½ ft. from the screen. A black box was attached to the screen and housed a fluorescent lamp that was 6 ft. long. A 5⁄16-in.-wide horizontal slit in the box formed an illuminated line, and segments of selected length (12, 36, and 60 in.) were exposed as standards.

Stimulus arrays, each containing six horizontal lines, were projected from negative transparencies onto the screen (see Fig. 1). Slides were designed to represent 17 stimulus values (mean lengths of the array) and covered a range of 22–183 cm. Lengths of lines projected from each slide were carefully calibrated with the projector aligned and securely mounted. Component lengths within arrays were selected in accordance with a standard rectangular distribution. The array having a mean length of 60 cm., for example, contained lines with measurements (in cm.) of 45, 50, 55, 65, 70, and 75. For all arrays the intercomponent range was 30 cm., and corresponding intervals between ordered components were approximately equal. Width of projected lines was ⅔ in. In each array the lines were randomly assigned to vertical positions 7½ in. apart and to lateral positions spaced at 1-cm. steps within a range 1.5 times the longest component. Stimuli were viewed binocularly through red goggles, and no additional cues (screen borders, walls, ceiling, or floor) were visible to *O*.

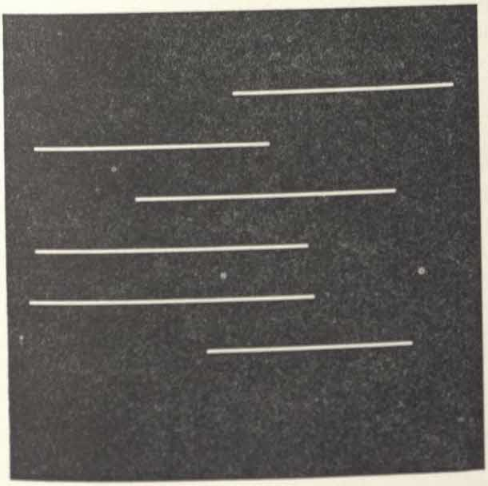


FIG. 1. Stimulus array with mean length of 95 cm.

Procedure. On arrival, *O* was given a simple test on arithmetic means and an orientation task in which he adjusted the length of a bar to indicate average length of several displays. He was then fitted with red goggles and escorted by flashlight into the darkened projection room. After further instructions, the 17 stimulus sets were each presented three times in an irregular order, a different sequence being followed for every *O*. Components of all sets were shown in two spatial arrangements. On every trial, stimulus arrays were exposed for 15 sec., followed by a 15-sec. dark period during which *O* announced his judgment. A pause of 1–2 min. was required after Judgment 35 to permit change of slide trays.

Five groups of *O*s were exposed to the same procedure. For three groups of 15 *O*s (composed of approximately equal numbers of men and women), an illuminated 'standard' line (12, 36, or 60 in. long) was also presented, 14–18 in. below each array, and the moduli were prescribed by *E* (10, 30, or 50 respectively). The selected standard was illuminated when *O* entered the projection room and remained on until the experiment ended.

In order to reduce constraints under which judgments were made, the standards for two additional groups were exposed before (and *only* before) the judgment series began. These standards consisted of parallel arrays. For this fourth group, of 20 *O*s (all men), the standard was an array with mean length of 37.4 in.; the *O*s were asked to call the average length '30.' In the fifth and last group, of 15 men and women, each *O* received a different standard array and selected the modulus. Instructions for all groups emphasized the importance of judging average length in proportion to the standard. This excerpt from instructions given the group receiving a standard 37.4-in. array should illustrate the procedure:

> "During the experiment pictures will be projected on the screen in front of you.
>
> "You'll notice there are six lines in the picture now on the screen. Every picture in the experiment will contain a group of six lines. Your job will be to look over the six lines in a group and judge how long they are on the average. You are not expected to calculate the average by arithmetic, as you would find the average of a column of numbers. Instead, look carefully at each line in the group and then estimate their average length as well as you can.
>
> "Rather than judge average length in inches or some other familiar unit, I'd like you to call the average length of the group you see now on the screen '30.' I'd also like you to use this picture as a standard for the rest of the experiment. For example, if the lines in another picture should appear twice as long on the average as the group now on the screen, then their average length would be two times 30, or 60. On the other hand, if the lines in another picture appeared, let us say, one-fourth as long on the average as the group now on the screen, then their average length would be one-fourth of 30, or 7.5. In other words, the average length of every group of lines presented in the experiment should be judged in proportion to the group of lines you now see on the screen, whose average length is called 30."

RESULTS AND DISCUSSION

To facilitate comparisons among groups, all judgments were transformed to a common scale of inches. Analysis of variance and median tests failed to disclose significant effects associated with *O* sex, and hence data of men and women were combined in further analyses. For every value of the variable stimulus, a geometric mean was calculated over transformed magnitude estimations of *O*s within each of the five treatments. Functions of the form

$$\Psi = kS^n ,$$

where Ψ is subjective magnitude, S is stimulus magnitude, and k and n are parameters, were fitted by the log-log reduction procedure to the geometric means of each group.[4] Parameters of these functions are presented in Table I and graphed in Figs. 2 and 3. It may be seen (Figs. 2 and 3) that the data cluster very closely around the fitted curves and appear to be described fairly well.

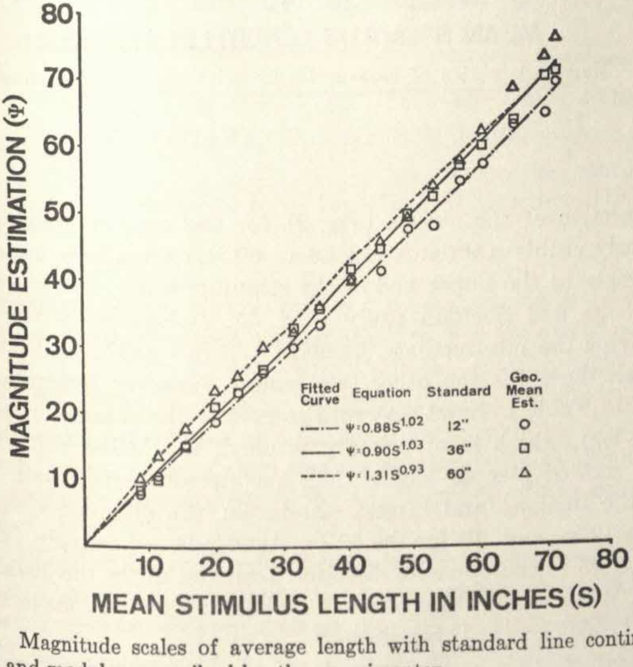


FIG. 2. Magnitude scales of average length with standard line continuously visible and modulus prescribed by the experimenter

---

[4] D. Lewis, *Quantitative Methods in Psychology,* 1960, 51–55.

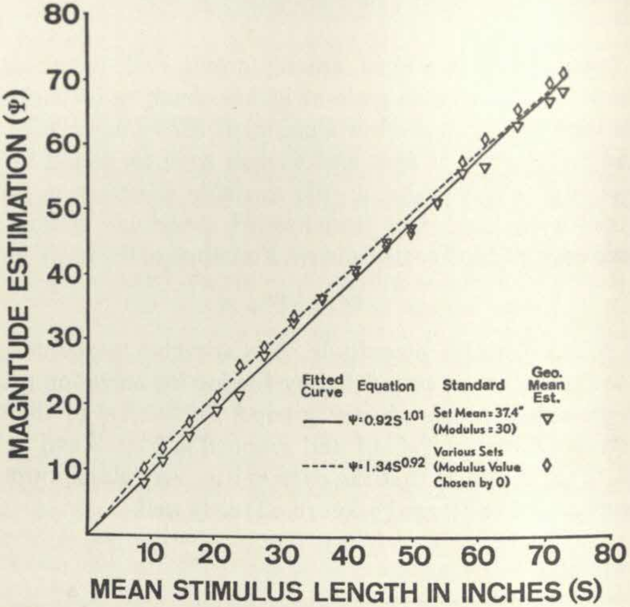


FIG. 3. Magnitude scales of average length with standard set exposed only at outset

Inspection of the curves (Fig. 2) for the groups receiving continuously visible standards (12, 36, or 60 in.) reveals no pronounced divergence at the upper end of the stimulus range, as there was in the Miller and Sheldon study (Fig. 4). Judgments of the group employing the intermediate standard (36 in.) again tended to fall between those for the other two groups, however (compare Figs. 2 and 4). Table I shows that with respect to the parameter n (slope in log-log), which is of primary interest, the Miller and Sheldon study showed steeper slopes than the present experiment, under both the smallest and largest standards: the differences were .12 for the 12-in. and .10 for the 60-in. standards respectively. A small shift of .03 in the opposite direction occurred under the 36-in. condition. While small, these effects may be due to variations in experimental conditions or *O* populations.

In both studies, a relatively steep slope was associated with the smallest standard (12 in.) and a flat slope with the largest (60 in.). Analysis of variance revealed significant differences among the

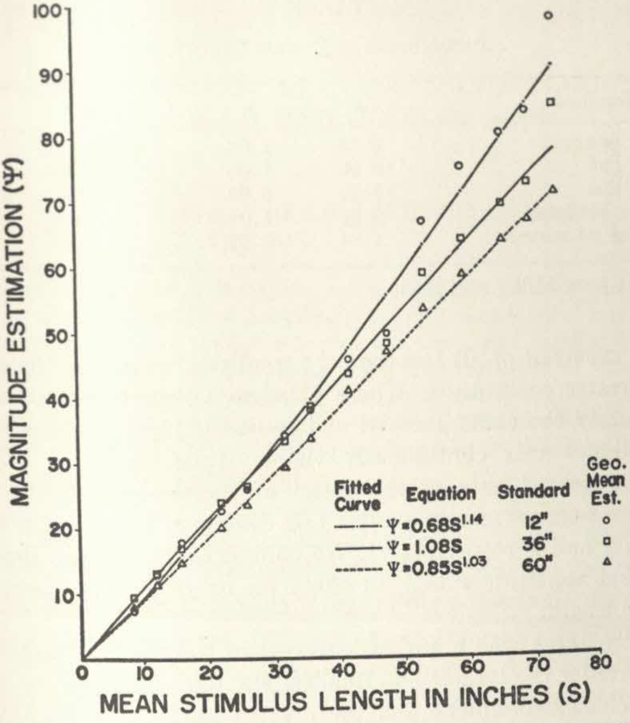


FIG. 4. Magnitude scales of average length with standard line continuously visible and modulus prescribed by the experimenter (data from Miller and Sheldon)

three standard groups of the Miller and Sheldon study [F (2, 15) = 5.90, $p < .025$] but not among the five groups of the present study [F (4, 75) = 1.20, $p > .10$]. While this finding might be attributed to differences between *O* populations or in experimental conditions, we cannot rule out the possibility that real differences between the groups receiving continuous standards in the present study may have been obscured by the presence of the other two groups.

The effects of lowered constraints may be ascertained by comparison of the first three groups of the present experiment with the last two. In the fourth group, where the 37.4-in. array was assigned a modulus of 30, the slope was only .02 less than the maximum slope under the conditions of greater constraints. Furthermore, the low-constraint fifth group, having individual standards and moduli,

TABLE I

PARAMETERS OF POWER FUNCTIONS

Standard	k	n	k†	n†
12-in. line	0.88	1.02	0.68	1.14
36-in. line	0.90	1.03	1.08	1.00
60-in. line	1.31	0.93	0.85	1.03
37.4-in. set mean	0.92	1.01		
Various set means	1.34	0.92		

† Data from Miller and Sheldon.

had an exponent of .01 less than the smallest exponent in the groups with greater constraints. Where physical values of the standards were nearly the same (second and fourth groups), yet one standard (36 in.) was continuously visible while the other (37.4-in. array) appeared only at the outset of the experiment, the fitted functions were extremely similar (.02 difference between the values of both n and k respectively). We cannot conclude from these results that constraints had an effect on magnitude estimations of average length.

Finally, there was a curious reversal in the relative positions of the curves for the largest and smallest standards between the Miller and Sheldon experiment and the present study (compare Figs. 2 and 4). While it may be tempting to give an interpretation to this finding (*e.g.* in terms of here darkened vs. there illuminated visual fields), a satisfactory explanation awaits further research.

In general, magnitude estimations of average length obtained under the two experiments' quite dissimilar observation conditions (different exposure durations, monocular vs. binocular vision, impoverished vs. highly illuminated visual field with visible contextual cues), with different observer populations (Army enlisted men vs. college men and women), and under different constraints on observers' judgments, yielded relationships between subjective magnitude of average length and stimulus magnitude that are described by power functions with slopes ranging only from .92 to 1.14 and tending to cluster around 1.00–1.03.

EFFECTS OF STRAIN, SEX, AND ILLUMINATION ON OPEN-FIELD BEHAVIOR OF RATS

FRED P. VALLE, *University of British Columbia*

Abstract. The effects of strain, sex, and level of illumination on open-field behavior were investigated with black-hooded and agouti-selfed rats. Results indicated that (*a*) black-hooded rats locomote more, rear equally often, and are less prone to leave the peripheral wall (more thigmotaxic) than agouti-selfed rats; (*b*) female rats locomote more, rear more, but show the same degree of thigmotaxis as males; and (*c*) increases in illumination produce decreases in both locomotion and rearing but increases in thigmotaxis.

On the basis of a wide variety of behavioral tests, Keeler concluded that laboratory strains of *Rattus norvegicus* which are homozygotic for the recessive 'black' allele at the A locus (*i.e.* rats which, if pigmented, have black fur) are tamer than agouti strains (*i.e.* rats which, if pigmented, have brownish-gray fur).[1] Broadhurst, however, reported that two black-hooded strains locomoted less in an open field than either an agouti-selfed strain (a nonhooded, agouti-colored strain) or an agouti-hooded strain.[2] If increased tameness is equated with increased locomotion, then Broadhurst's data suggest either that Keeler's hypothesis is incorrect or, at the very least, that the presumed taming effects of the black allele can be modified by the genetic context in which it occurs.

The present study was an attempt to shed additional light on the behavior of black and agouti strains in an open field, measuring not only the amount of locomotion and amount of defecation but also the frequency of rearing and degree of thigmotaxis, or wall-hugging behavior. If Keeler's hypothesis is correct, the black strain should locomote more, rear more, defecate less, and be less

---

Received for publication July 27, 1969. The author thanks Michael Judd, who helped collect the data, and the President's Research Committee of the University of British Columbia, which provided financial assistance.

[1] C. E. Keeler, The association of the black (nonagouti) gene with behavior in the Norway rat, *J. Hered.*, 33, 1942, 371–384.

[2] P. L. Broadhurst, Determinants of emotionality in the rat: III, Strain differences, *J. comp. physiol. Psychol.*, 51, 1958, 55–59.

thigmotaxic (on the assumption that wall-hugging reflects, in part, the effects of timidity) than the agouti strain.

In addition to the strain variable, the study also investigated the effects of sex and level of illumination on the above open-field measures. Previous studies, reporting only the locomotion measure, suggest that females should be more active than males[3] and that rats should be more active in the dark than in the light.[4] The present study, however, is apparently the first attempt to assess the effects of strain, sex, or level of illumination on either rearing behavior or wall-hugging behavior.

METHOD

Summary of design. Two levels of illumination, two strains of rats, and the two sexes of each strain, formed a 2 × 2 × 2 between-groups factorial design with measures on seven subjects within each cell. Hence, the total $N = 56$.

Subjects. The 28 BD-IX rats were bred from stock generously supplied by Dr. D. Müller-Schwartz of San Francisco State College. The BD-IX strain[5] is homozygotic for the dominant (agouti) alleles at the A locus and for the dominant (selfed, or nonhooded) alleles at the H locus. In addition, this strain is homozygotic for the dominant alleles at the C locus (rats which are homozygotic for the recessive alleles at this locus are albino) and for the dominant alleles at the P locus (rats which are homozygotic for the recessive alleles at this locus have golden fur and pink eyes).[6]

The 28 Long Evans rats were bred from stock originally obtained from National Farms, Edmonton, Canada. This strain is homozygotic for both the recessive *aa* (black) alleles and the recessive *hh* (hooded) alleles; see n. 1. In addition, since the Long Evans rats are pigmented, the strain must either be heterozygotic at the C locus or homozygotic for the dominant alleles, as is the BD-IX strain. Hence, at the very least, the two strains differ genetically at two loci; they may also differ at other loci as well, although such differences are not cataloged.

Apparatus. The open field was a square white board, 1.22 m. on a side, divided by black lines into 16 equal squares. It rested on the floor on .05-m. legs and had a .30-m. wall around the periphery. It was located in a windowless room that was 3.05 m. long, 2.64 m. wide, and 2.94 m. high. During observation, illumination was provided by either a 150-w. clear white bulb (bright condition) or a 25-w. G.E. Ruby Red light (dark condition) suspended from the ceiling, 2.76 m. above the center of the field.

---

[3] R. N. Hughes, Effects of food deprivation, deprivation experience, and sex on exploration in rats, *Brit. J. Psychol.*, 59, 1968, 47–53; R. N. Hughes, Behaviour of male and female rats with free choice of two environments differing in novelty, *Anim. Behav.*, 16, 1968, 92–96.

[4] A. P. Mead, A quantitative method for the analysis of exploratory behaviour in the rat, *Anim. Behav.*, 8, 1960, 19–31; W. I. Welker, Escape, exploratory, and food-seeking responses of rats in a novel situation, *J. comp. physiol. Psychol.*, 52, 1959, 106–111.

[5] Von H. Druckrey, P. Danneberg, W. Dischler, and D. Steinhoff, Reinzucht von 10 Rattenstämmen (BD-Stämme) und Analyse des genetischen Pigmentierungs-Systems, *Arzneim-Forsch.*, 12, 1962, 911–919.

[6] R. Robinson, *The Genetics of the Norway Rat*, 1965.

Procedure. All the rats were born in the laboratory and reared under 24-hr. artificial illumination. Each rat was handled every four or five days up to weaning. Approximately four weeks after weaning, the animals were separated into like-sex group cages. They were not handled again until one month before testing, when they were moved into individual cages and given five brief handlings each week for four weeks. Each rat was then tested once between 5:30 and 7:30 P.M. when approximately 150 days old.

On the test day, after random assignment (within litters) to one of the two illumination conditions, the rats were transported singly to the test room where they remained alone for 8 min. to adapt to the lighting condition which would prevail during the test. *E* then reentered the room and 2 min. later placed the rat (by the tail) into the corner of the field closest to the point where *E* stood while observing it. The rat was allowed to explore for 10 min. with records kept of (*a*) number and pattern of squares entered, with 'entrance' defined as placing both forepaws in a square, (*b*) number of seconds spent in the four central squares, (*c*) number of rears, and (*d*) number of boli. After each rat was tested, the field was washed with a mild vinegar solution and dried. The next rat was then brought into the room for the 10-min. adaptation period, and so fourth. To minimize any odor effects not removed by the washing procedure, the males of a strain were tested one week and the females the following week. There was a two-month interval between strains.

RESULTS

Fig. 1 shows the effects of sex, illumination level, and strain on total number of units entered during the 10-min. test. Analysis of variance indicated that all three differences shown in Fig. 1 were significant: females locomoted more than males ($F = 5.05$, $df = 1, 48$, $p < .05$); animals entered more squares in the dark than in the light [$F = 23.41$, $df = (1, 48)$, $p < .001$]; and the Long Evans rats locomoted more than the BD-IX rats [$F = 13.46$, $df = (1, 48)$, $p < .001$]. There were no significant interactions.

Fig. 2 shows the effects of sex, illumination, and strain on number of rearings. It can be seen that females reared more than males, that animals reared more in the dark than in the light, and that there was virtually no difference in frequency of rearing between the two strains. Analysis of variance indicated the effects of sex [$F = 16.58$, $df = (1, 48)$, $p < .001$] and illumination [$F = 43.19$, $df = (1, 48)$, $p < .001$] to be significant. There were no reliable interactions.

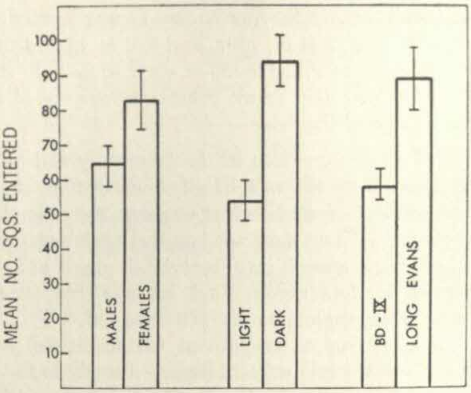


FIG. 1. Effects of sex, strain, and level of illumination on mean number of squares entered by rats during a 10-min. open-field test. The effects of each independent variable are shown collapsed over the other two. The vertical lines represent the standard error of the mean.

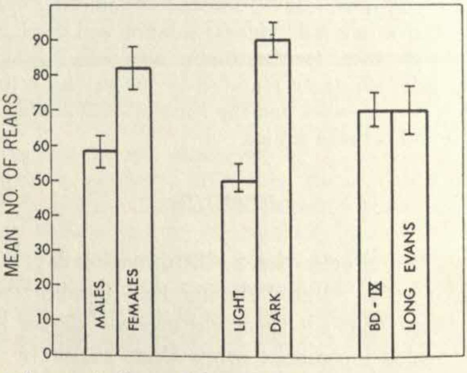


FIG. 2. Effects of sex, strain, and level of illumination on mean number of rears by rats during a 10-min. open-field test. The effects of each independent variable are shown collapsed over the other two. The vertical lines represent the standard error of the mean.

Fig. 3 illustrates the effects of the three independent variables on number of seconds spent in the central area of the open field. Medians are used because one extremely deviant score in the group of male BD-IX rats tested under bright illumination markedly inflated the mean of the group and greatly increased the variance. It can be seen from Fig. 3 that females spent somewhat more time in the central area than did males, that rats in the dark spent almost

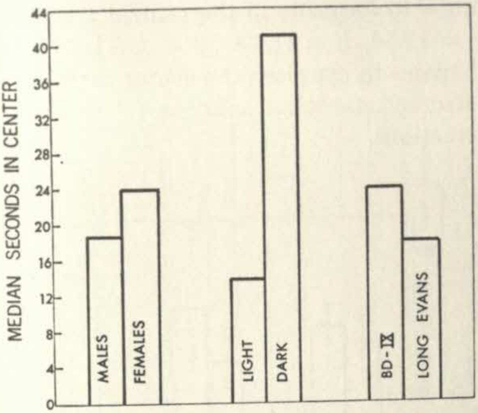


FIG. 3. Effects of sex, strain, and level of illumination on median number of seconds spent in the central area of an open field by rats during a 10-min. test. The effects of each independent variable are shown collapsed over the other two.

three times as many seconds in the central area as did animals in the light, and that BD-IX rats spent somewhat more time in the central area than did Long Evans rats. Kruskal-Wallis analyses of ranked data indicated that only the illumination variable produced a significant effect: $H = 14.36$, $df = 1$, $p < .005$.

Fig. 4 illustrates the effects of sex, illumination and strain on the 'thigmotaxic ratio,' *i.e.* the number of central squares entered divided by the total number of squares entered. It should be noted that if there were no thigmotaxic factor operating, *i.e.* if rats were entering squares randomly with respect to location, the expected value of the thigmotaxic ratio would be .25, since the central four squares of the field constitute 25% of the total area. Given this expected value, all groups except the male BD-IXs tested under bright illumination could be said to be thigmotaxic, since all had mean thigmotaxic ratios which were at least two standard errors less than .25. The exceptional group, it will be recalled, contained the animal which spent most of the test period in the central area of the field. This unusual animal also had the highest thigmotaxic ratio of the study: .37. Despite the peculiarities of this animal and the variance it introduced, analysis of variance supported the pattern of differences illustrated in Fig. 4; namely, animals tended to locomote in the central area relatively more often in the dark than in the light [$F = 4.28$, $df = (1, 48)$, $p < .05$] and BD-IX animals were relatively more prone to locomote in the central area than were Long Evans rats [$F = 18.56$, $df = (1, 48)$, $p < .001$]. The slightly greater tendency for females to approach the center as shown in Fig. 4 was not substantiated by statistical analysis ($F < 1$). There were no significant interactions.

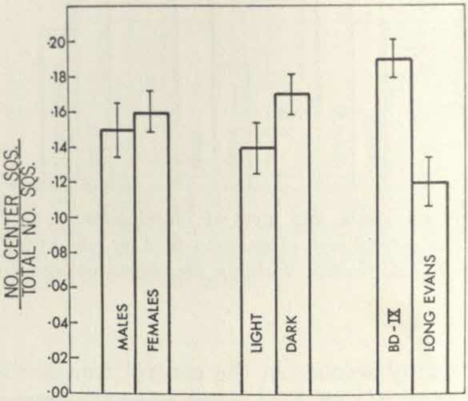


FIG. 4. Effects of sex, strain, and level of illumination on mean ratio of central squares entered to total squares entered by rats during a 10-min. open-field test. The effects of each independent variable are shown collapsed over the other two. The vertical lines represent the standard error of the mean.

Fig. 5 shows the effects of the three independent variables on number of boli excreted during the test. Females defecated more than males, BD-IXs defecated more than Long Evanses, and rats in the light defecated more than rats in the dark. None of these differences was statistically reliable, however: F (strain) $= 2.01$, $df = 1, 48$; F (sex) < 1; F (illumination) < 1. There were no significant interactions.

Table I summarizes the effects of the three independent variances on each of the five dependent variables discussed.

DISCUSSION

The results of this study permit neither clear confirmation nor clear rejection of Keeler's hypothesis (see n. 1) that rats homozygotic for the black allele are tamer than their agouti counterparts. On the one hand, the finding that, contrary to Broadhurst's

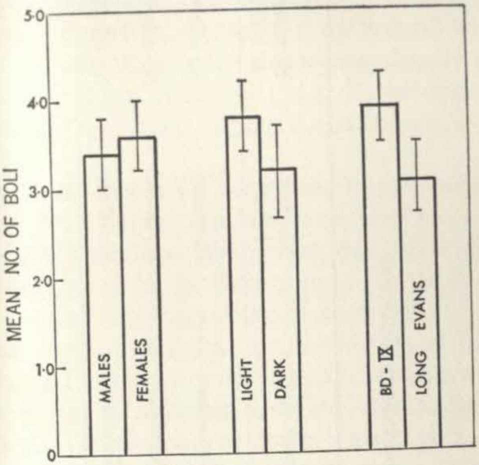


FIG. 5. Effects of sex, strain, and level of illumination on mean number of boli excreted by rats during a 10-min. open-field test. The effects of each independent variable are shown collapsed over the other two. The vertical lines represent the standard error of the mean.

report (see n. 2), black rats locomoted significantly more than agouti rats supports Keeler's conclusion, if it is assumed that increased locomotion represents increased tameness. On the other hand, if it is assumed that decreased thigmotaxis (or wall-hugging) also represents increased tameness, then the present results refute Keeler's hypothesis, since the agouti strain was relatively more prone to leave the wall and approach the center of the field than was the black strain. Finally, the lack of a strain effect on the otherwise-sensitive rearing measure suggests that these three kinetic aspects of exploration—locomotion, wall-hugging, and rearing—are not equally sensitive, or even sensitive in the same manner, to genetic differences between the two strains.

An additional complexity is encountered when the defecation data are considered: the lack of a reliable strain difference suggests that the two strains did not differ in terms of 'emotionality' in the test situation. Since, however, the defecation data also failed to yield a reliable effect for either the sex or illumination variable, it may be that defecation is simply an insensitive measure with well-handled rats.

The effect of sex on open-field behavior is consistent with that reported by other investigators: the females locomoted more than

TABLE I

Effects of Three Independent Variables on Five Dependent Measures of Open-Field Behavior in Rats

Independent variable	Total number of squares entered	Number central squares / Total number squares	Amount of time in central area	Number of rears	Number of boli
Illumination					
Darkness	Increases	Increases	Increases	Increases	No effect
Sex					
Female	Increases	No effect	No effect	Increases	No effect
Strain					
BD-IX	Decreases	Increases	No effect	No effect	No effect

males.[7] The present data also show that females reared significantly more than males. Interestingly enough, there was no sex difference in the thigmotaxic data. Hence, females were generally more active than males but the pattern of their activity did not differ from that of males, at least insofar as the thigmotaxic ratio reflects the pattern of locomotion.

The effect of illumination of locomotion is also consistent with that reported in other studies: the rats locomoted more in the dark than in the light.[8] In addition, the present data show that the rats reared significantly more in the dark than in the light. It is also of interest that the rats were significantly more prone to leave the wall in the dark than in the light. One implication of this finding is that, in addition to the tendency to maintain contact with a lateral surface,[9] the tendency to avoid open, lighted areas also contributes to wall-hugging, or thigmotaxic, behavior. This in turn is consistent with the assumption (made above) that decreases in wall-hugging represent decreases in fearfulness.

---

[7] Hughes (n. 3 above, second reference).

[8] Welker (n. 4).

[9] W. J. Crozier and G. Pincus, Stereotropism in rats and mice, *J. gen. Physiol.*, 10, 1926, 195–203.

EFFECT OF A MOVING VISUAL ENVIRONMENT ON LOCALIZATION OF SOUND

WILLARD R. THURLOW and THOMAS P. KERR, *University of Wisconsin*

Abstract. Rotation of a striped visual environment about a subject caused displacement, in the direction of that rotation, of the perceived location of (1) the 'straight ahead' (of the head) and of (2) a source of sound. The mean shift in the location of the sound was consistently greater than the mean shift in the location of the 'straight ahead.' It is suggested that these effects are related to the nystagmic eye movements occurring as a reaction to the moving environment.

Previous experiments on the effect of a moving visual environment on the displacement of a source of sound in the left-right direction have not had consistent results. In these experiments, the subjects have been placed at the center of a striped cylinder and required to localize sounds while the cylinder was either moving or stationary. Gemelli found that whether a displacement effect was obtained or not depended on whether the subject perceived himself as moving.[1] Arnoult did not obtain significant displacement effects.[2]

The present experiments were performed to investigate the situation further. Preliminary observation indicated that there were experimental conditions under which rotating the visual environment consistently shifted the localization of sound and that the shift in that localization was correlated with a shift in the perception of 'straight ahead.' It is certainly reasonable to expect a correlation between the perceived direction of the head and the location of a sound, since cues to head position must be taken into

---

Received for publication August 11, 1969. The research was supported by a grant from the National Science Foundation and facilitated by a research leave granted to the first author by the Graduate Research Committee of the University of Wisconsin. Both authors thank Richard C. Rollefson for his assistance in gathering the data of Experiments I, II, and III.

[1] A. Gemelli, The effect of illusory perception of movement on sound localization, in *Essays in Psychology,* dedicated to David Katz, 1951, 104–117.

[2] M. D. Arnoult, Localization of sound during rotation of the visual environment, this JOURNAL, 65, 1952, 48–58.

account in locating sounds. Indeed, McLaughlin and Bower found a correlation between the shift in auditory localization and the shift in 'straight ahead' in experiments with prism displacement of the visual field.[3] It was of interest to see whether displacement effects in the present experimental situation could be explained completely by shifts in the direction of 'straight ahead.'

EXPERIMENTS I, II, AND III

Method

Experiment I. The *S* was seated in a chair, with his head fixed in position by a bite bar. A vertical cylinder of striped sheeting 2.5 ft. in radius rotated around him at a rate of 1 revolution in 7 sec. The sheeting was commercially available striped sheeting (Fig. 1): the width of the darkest wide stripe (a dark blue) was 11 mm.; the brightness of the lightest stripe, under the overhead illumination used, was 17.6 mL., and the brightness of the darkest wide stripe was 3.1 mL. A direction-indicator rod could be rotated by *S*. Fig. 2 shows a schematic cross-section of the apparatus from the left side. The X indicates the position of the *S*'s ears as he sat in the chair.

A loudspeaker (Electro-Voice MC 8) was placed outside the cloth cylinder at ear level (on a stand) and 4.5 ft. from the center of the cylinder's axis (*i.e.* 4.5 ft. from a point midway between *S*'s ears as he sat in the apparatus). This loudspeaker was placed 32° to the right of straight ahead of *S*. It was not visible to *S* from inside the cloth cylinder. The sound was a high-frequency thermal noise. It was produced by passing thermal noise from a Grason-Stadler noise generator (Model 455C) through a Krohn-Hite filter (Model 330M) set to 7,500–8,000 band-pass. The noise signal then went through a Grason-Stadler electronic switch (Model 829-D), amplifier, and transformer, to the loudspeaker. The electronic switch was triggered by a Hunter timer to turn on the noise for a 10-sec. interval. At *S*'s ears the sound level (Scale A on the sound-level meter) was 51 db. The experiments were carried out in a room with fiberglass blanket on the walls to minimize reflections from the walls.

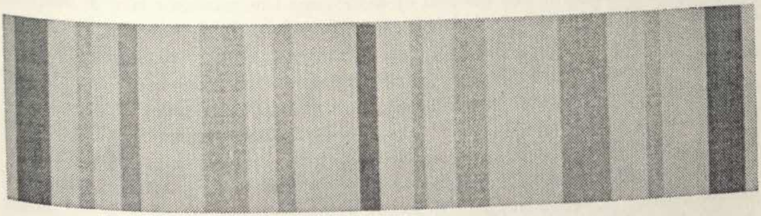


FIG. 1. Photograph of vertical stripes of cloth cylinder

---

[3] S. C. McLaughlin and J. L. Bower, Auditory localization and judgments of straight ahead during adaptation to prism, *Psychon. Sci.*, 2, 1965, 283–284.

The *S* first had practice in (*a*) lining up the indicator rod to 'straight ahead' and (*b*) lining up the indicator rod with the direction of the sound. *S*s wore goggles of transparent plastic, with masking tape applied to them in such a way that the *S*s could not see their bodies or other fixed reference points above or below them.

The experimental observations made with the striped environment moving were preceded and followed by control observations made with the striped environment stationary. In the control observations, the *S* (*a*) tried to line up the vertical rod with what appeared to him to be straight ahead. Then (*b*) he tried to line up the indicator rod with the perceived direction of the sound's source. Then the observations were repeated in the reverse order. In the experimental observations, the *S* was instructed to look at the stripes on the sheet (not at the rod) and try to make the stripes 'stand still' perceptually, and to imagine himself as rotating while he made observations of the 'straight ahead' and then of the sound. This procedure was then repeated with observations in the reverse order. In all cases where the environment was moved, rotation was to the left.

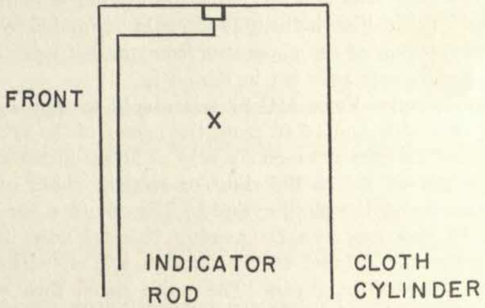


FIG. 2. Schematic of experimental setup. The X marks the position of the subject's ears.

Experiment II. With two exceptions, this experiment was run in the same way as Experiment I. (1) Instead of *S* moving the indicator rod, *E* moved it and *S* judged when it was in the proper location. (This change in mode of judgment was made to check the possibility that results of Experiment I might have been related to an effect on *S*'s motor response.) Separate judgments were made when the indicator was started from the left and when it was started from the right. There were, therefore, twice as many judgments in Experiment II as in Experiment I. (2) The loudspeaker was moved to a position 5° to the left of straight ahead of *S*. This was done in order to have the loudspeaker position closer to (but not identical to) being straight ahead of *S* for further comparisons of his judgments of location of sound vs. 'straight ahead.'

Experiment III. Experiment III was run in the same way as Experiment II, except that *S* was additionally instructed to define 'straight ahead' as "the direction in which your head is pointing." This was done on the grounds that the term 'straight ahead' may be ambiguous if perceived head direction is not the same as perceived body direction.

Results

Average differences between the results when the environment was moving and when it was stationary are shown in Table I. It is apparent that there was a shift both in the localization of the sound and in the position of 'straight ahead.' This shift was to the left—in the direction of movement of the striped environment. All mean shifts were statistically significant, beyond the .01 level (by standard *t* tests). Very few of the *S*s failed to show the shifts. All *S*s reported that they perceived themselves as moving, in a direction opposite to the direction of movement of the striped environment. The average amount of shift in localization of sound was consistently greater than the average amount of shift in 'straight ahead' by 3°–5°, although the difference reached statistical significance only for Experiment II. In this case, the difference was significant beyond the .01 level.

EXPERIMENTS IV, V, AND VI

Method

Experiment IV. This experiment was run to find whether perception of self-movement was necessary for the displacement effects. Exploratory experimentation showed that the perception was most likely to arise when the vertical indicator rod was in the field of view and less likely to arise if the time interval of exposure to the moving stripes was short. Therefore, in this experiment, a nonvisual indicator response was used, and the time of exposure was shortened by having *S* close his eyes except when making judgments.

With several exceptions, the procedures were the same as in Experiment III. (1) The vertical rod was *not* here used for the indicator response. *S* pointed with his index finger (arm extended) to indicate (*a*) direction of "straight in front of his head," (*b*) direction of the source of sound. *S*'s pointing hand was not in his field of view. He was not told to try to imagine himself as moving. (2) *S* was told explicitly that the sound was coming from a position at head level, in front. It was found necessary to introduce this instruction in order to minimize cases of localization of the sound above or behind the *S*. The source of sound was actually put 5° to the right of straight ahead of *S* (in terms of his head position). (3) *S* kept his eyes closed except when making judgments. Thus, the duration of exposure to the rotating stripes was very close to 10 sec. (4) The number of judgments made by each *S* was the same as in Experiment I.

Experiment V. Because of the rather small displacement effects in Experiment IV, Experiment V made two changes aimed at decreasing variability in average results, in order to get increased power of statistical tests. (1) More than twice as many *S*s were run. (2) A wider band of noise, 6,000–10,000 cps, was used, at 52 db., Scale A. One minor change in procedure was also introduced: (3) *S* opened his eyes each time before he was given instructions as to which location judgment he was to make.

Experiment VI. This experiment was run essentially like Experiment V, except that the indicator response was changed back to that of Experiment II and *S*s were asked to try to imagine themselves as moving. We wanted to find whether we would get results comparable to those obtained in Experiments I, II, and III if we returned to the procedures which had there produced perceived self-movement. In this experiment, then, (1) the vertical indicator rod was used, moved by *E* as in Experiment II; readings were taken with the vertical rod brought in from the left only. (2) *S*s were asked to try to imagine themselves as moving (for the moving-stripe condition). When they felt themselves rotating, location judgments were measured.

Results

Experiment IV. Nine *S*s of 10 tested did not perceive themselves moving. The average results for these nine *S*s are shown in Table I. The mean shift in localization of the sound and the difference between that shift and the mean shift in position of 'straight ahead' were both statistically significant beyond the .05 level. The shift in perceived position of 'straight ahead' was not statistically significant. It is of interest that the sizes of displacements of the sound and of the 'straight ahead' were smaller than in Experiments I, II, and III. However, the difference (4.6°) between the mean shift in localization of the sound and the mean shift in 'straight ahead' was of the same order of magnitude as that obtained in Experiments I, II, and III. Apparently, then, perception of self-movement is not necessary for production of a displacement in localization of sound.

Experiment V. There were 20 *S*s in Experiment V (out of 28 run) who did not perceive themselves as moving. These 20 *S*s gave mean shifts of 4.2° and 2.7° for localization of the sound and for 'straight ahead' respectively (see Table I). The *t* tests showed that both of these mean shifts were significantly different from zero, and were significantly different from each other (beyond the .05 level).

Experiment VI. Results for the 10 *S*s tested are given in Table I. Mean shifts in localization of sound and in 'straight ahead' were significant beyond the .01 level. The mean difference between the two shifts was significant at the .05 level. The results seem quite comparable with those obtained in Experiments I, II, and III, which were also run with the vertical rod as an indicator of the localization of sound and of the 'straight ahead.' Further *t* tests showed that the magnitude of mean shifts in localization of sound and in 'straight ahead' was significantly greater in Experiment VI than in Experiment V (beyond the .05 level).

TABLE I

MEAN SHIFT IN DEGREES FOR LOCALIZATION OF SOUND AND FOR 'STRAIGHT AHEAD' AS A RESULT OF MOVEMENT OF STRIPED ENVIRONMENT

	Experiments					
	I	II	III	IV	V	VI
Number of subjects	13	10	10	9	20	10
Perception of self-movement	Yes	Yes	Yes	No	No	Yes
Shift in localization of sound (a)	11.7**	10.2**	12.9**	5.1*	4.2**	12.4**
Shift in 'straight ahead' (b)	8.5**	5.3**	9.2**	.5	2.7*	7.1**
Difference (a − b)	3.2	4.9**	3.7	4.6*	1.5*	5.3*

* Significant at .05 level.
** Significant at .01 level.

DISCUSSION

The results showed that a moving, vertically striped environment has a consistent effect both on localization of sound and on location of 'straight ahead' (of the head). In both cases there was a displacement effect in the direction of rotation of the environment, but the displacement was typically significantly greater for localization of sound than for location of 'straight ahead.' Perception of the self as moving was not necessary for the displacement effects. Even though displacement effects were found to be larger in those experiments where the subjects perceived themselves as moving, we are not able to attribute the increased displacement effects to perceived movement because of other variables involved. For example, the increased displacement might have been due to a longer exposure time to the moving stripes, or to use of a visual indicator for localization judgments. Further experiments are necessary to separate these variables.

It seems unlikely that the auditory displacement effect is related simply to degree of *perceived* movement of the visual environment. We tested seven more subjects in the same manner as in Experiment V but had them fixate the end of a small rod, 14 in. in front of their eyes. These subjects showed an average auditory displacement effect of only 2°. When the eyes are fixated in this way, the stripes appear to move much faster then when one looks at the stripes and tries to make them 'stand still.' If auditory displacement effects were related to perceived visual movement, one would expect a much larger displacement effect when the fixation point was used. However, this was not the case.

It seems more likely that the displacement effects are mainly related to nystagmic movements of the eyes. It is well known that nystagmic movements occur to moving stripes.[4] The eyes move with the stripes, then return quickly and start a new cycle. The average position of the eyes would then be, not physically straight ahead of the head, but in the direction of rotation of the stripes. It is also known that if the eyes are shifted in direction, the perceived position of 'straight ahead' is shifted in that same direction.[5] Thus we may have a basis for understanding why the position of 'straight ahead' is shifted by rotation of the visual environment.

If the position of 'straight ahead of the head' is shifted, it is reasonable to suppose that the localization of sound would be shifted —since that localization must take cues of head position into account. But there is an additional consistent shift in the localization beyond that to be expected from the shift in 'straight ahead.' It is conceivable that there may be an additional, direct biasing effect of eye position on the sound-localization system.

[4] I. P. Howard and W. B. Templeton, *Human Spatial Orientation*, 1966, 55-62.
[5] Howard and Templeton (see n. 4 above), 288.

EFFECTS OF REWARD STRUCTURE AND COGNITIVE DIFFERENCE IN A MIXED-MOTIVE TWO-PERSON CONFLICT SITUATION

Leon Rappoport, *Kansas State University*
and George Cvetkovich, *Western Washington State College*

Abstract. A two-factor reward structure permitting variation of the degree to which one person may profit at the other's expense was employed to study conflict between persons working together on an uncertain judgment task. The results show that conflict varied as a complex function of reward factors, experience with the task, and discrepancy between the persons' initial views of the task. Judgmental accuracy was also affected. These findings suggest that by emphasizing unitary reward structures at the expense of important psychological factors, prior studies of mixed-motive conflicts have oversimplified a complex phenomenon.

The literature on interpersonal conflict (IPC) is dominated by gaming formulations which treat competitive reward structures as the cause of conflict. In zero-sum games, a player can only profit at the expense of his opponent; in nonzero-sum games, the same condition obtains, but here both players can make an intermediate profit if neither tries for the maximum.[1] An alternative approach to IPC treats discrepant meanings as the cause of conflict.[2] And it has been shown that persons attributing different meanings to a given cue disagree even in the presence of a fully cooperative reward structure.[3] Heretofore, however, mixed-motive conflicts have

---

Received for publication March 27, 1969. This work was supported by NSF Grant GS-1912. Partial results were reported at the 1969 convention of the Southwestern Psychological Association. The authors thank Terry Applegate and Lodis Rhodes for their help in running the experiment.

[1] M. Deutsch and R. Krauss, Studies of interpersonal bargaining, *J. Confl. Resol.*, 6, 1962, 52–76; A. Rapoport and C. Orwant, Experimental games: A review, *Behav. Sci.*, 7, 1962, 1–37.

[2] K. Hammond, New directions in research on conflict resolution, *J. soc. Issues*, 21, 1965, 44–66.

[3] L. Rappoport, Interpersonal conflict in cooperative and uncertain situations, *J. exp. soc. Psychol.*, 1, 1965, 323–333; K. Hammond, F. Todd, M. Wilkins, and T. Mitchell, Cognitive conflict between persons: Application of the "lens model" paradigm, *J. exp. soc. Psychol.*, 2, 1967, 343–360.

been studied only according to the model for nonzero-sum games. The present study investigated them according to a modified version of Hammond's cognitive-conflict paradigm.

METHOD

In its original form, the cognitive-conflict paradigm involved "two persons who (1) attempt to solve problems which concern both of them, (2) have mutual utilities (their gain or loss derives from their approximations to the solution of the problem), (3) receive different training in the solution of a problem involving uncertain inference, (4) are then brought together and find themselves dealing with a familiar problem which their experience apparently prepared them for, but (5) find that their answers differ, and that neither is as good as it has been, although each answer is logically defensible and, (6) who must provide a joint decision as to the correct solution and, therefore, (7) must adapt to one another as well as to the task if they are to solve their problems. These conditions were chosen so that the research paradigm would include those circumstances in which men find themselves when attempting to deal with problems to which they bring different and inadequate cognitive systems."[4]

In the present experiment, point (2) above was changed, to give *Ss* *related* utilities: their individual gain depended both on their collective approximation to the correct solution and on the degree to which each could impose his own viewpoint on the other as they worked to agree on a joint answer. The reward structure can be expressed symbolically as a two-factor equation:

$$IG = f(CA) + f(P) ,$$

where IG is individual gain, CA is collective accuracy, and P is persuasion of the other. By differentially weighing CA and P, it was possible to specify the degree to which a competitive element was introduced into the experimental situation.

A second change from the original description of the paradigm concerns point (3) above. Instead of training *Ss* to interpret cues differently, they were here selected by means of a preliminary assessment instrument and paired according to whether they had different or similar views of the experimental problem. This selection procedure has been found to be as efficient as training when a realistic task is employed,[5] and it is desirable on other methodological grounds.

Task. The experimental task was to judge racial strife in 15 hypothetical communities. For each community, three 10-point cues described the degrees of integration in education, in housing, and in jobs respectively. Depending upon the values defined on these cues, the *Ss* were to agree on the amount of racial strife likely in each community. This criterion had 20 points, ranging from minimum to maximum strife, and included plausible descriptions of a realistic range of conditions of such strife. The mathematical structure of the 15-trial task was such that the housing cue correlated −.92 with the criterion, while the education and job cues were not significantly related to the criterion. None of the three cues was significantly correlated with the other two cues, and the multiple correlation between all three cues and the criterion was .93. The task therefore contained irreducible uncertainty, but since *S*s were allowed free discussion on every trial and were given feedback on their accuracy after each joint judgment, they learned to make better than chance judgments as trials proceeded.

---

[4] Hammond, Todd, Wilkins, and Mitchell (see n. 3 above), pp. 343–344.

[5] L. Rappoport, Cognitive conflict as a function of socially-induced cognitive differences, *J. Confl. Resol.*, 13, 1969, 143–148.

Design. The three experimental conditions manipulated as independent variables were these. (1) *IG ratios:* *S*s were paid after each trial according to either an 80CA/20P or a 20CA/80P ratio. This was calculated from a matrix explained in advance and kept before the *S*s throughout the experiment (see Fig. 1). Thus in the 80CA/20P condition, each *S* received 80% of the maximum payment possible on each trial depending upon how close the joint judgment was to the correct answer. He (*e.g.* S_1) received 20% of the maximum possible depending upon how far he could persuade the other (S_2) to depart from his initial judgment of the criterion in the direction of his (S_1's) initial judgment. The scheme was reversed in the 20CA/80P condition. (2) *IG levels:* in the low-pay condition, *S*s could make up to 30 cents on each trial. In the high-pay condition, they could make up to a dollar on each trial. (3) *Cognitive differences:* 80 college students, men selected according to their judgments of the amounts of racial strife indicated by rectangularly distributed test patterns of the three cues (*i.e.* the cognitive-difference assessment instrument) were arranged to provide 20 cognitively different (CD) and 20 cognitively similar (CS) pairs. These *S*s were drawn from a larger pool of approximately 150 who had all completed the assessment instrument. The CD pairs were composed of *S*s whose judgments on this instrument correlated an average of −.77 with one another; the average intercorrelation of those making up the CS pairs was +.79. For purposes of analysis, the 15 trials were treated as a fourth factor involving five three-trial blocks, thereby casting the entire experiment into the form of a four-way analysis-of-variance design with five pairs of *S*s in each cell.

RESULTS AND DISCUSSION

Conflict. Conflict was measured on every trial by the discrepancy between the initial judgments of each pair of *S*s as they sought to agree on a joint judgment. For the main effects, the results of the four-way ANOVA on conflict scores were: trials ($F = 5.75$), $p < .001$; IG levels ($F = 3.40$), $p < .10$; cognitive differences ($F = 3.18$), $p < .10$; and IG ratios ($F < 1.0$).[6] However, the most important results concern the interactions plotted in Fig. 2. The interaction of trials, IG level, and cognitive difference [($F = 4.16$), $p <$

---

[6] Full ANOVA summary tables are available upon request to the first author.

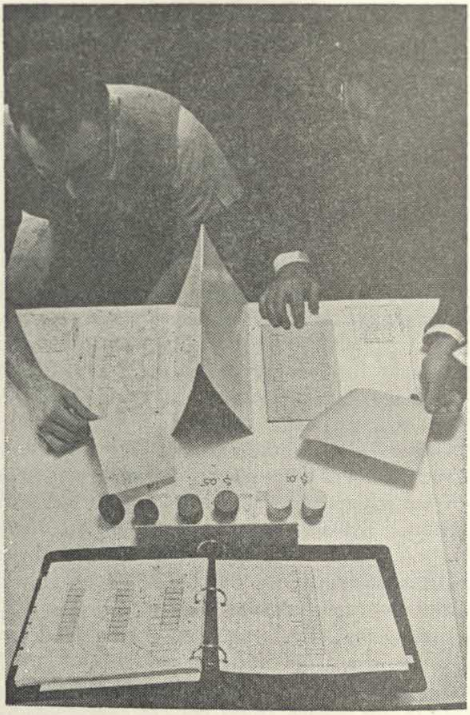


FIG. 1. Paired subjects working on experimental task

.01] indicates, as has been shown in prior studies of cognitive conflict (see nn. 4 and 5), that CD pairs have more conflict at the outset than CS pairs and that conflict is reduced in both groups with increasing trials. But IG level here interfered: instead of coming together over trials, the CD and CS pairs in the high-pay condition (high IG level) spread apart. Moreover, individual mean comparisons show that while CS pairs ended at about the same low level of conflict as that at which they began, CD pairs in the high-pay condition (high IG level) had significantly *more* conflict at the end than they did at the beginning. It seems clear that different pay levels have strikingly different effects on the course of conflict between *S*s who initially think differently about a mutual problem.

In a somewhat different way, the same conclusion applies to the effect of IG *ratios* on conflict. The overall interaction of trials, IG ratio, and cognitive difference [($F = 2.07$), $p < .10$] was only of borderline significance because of variability in the data following the first block of three trials. In Trial Block 1, however, data plotted (and individual mean comparisons) show that while conflict was similar for CD and CS pairs in the 80CA condition, it was sharply different for the pairs in the 80P condition. At both IG levels (high- and low-pay conditions), the 80P condition increased conflict in CD pairs; it had no significant effect for CS pairs in the high-pay conditions and a barely significant negative effect for CS pairs in the low-pay conditions. Note also that across all trial blocks in the high-pay conditions, the IG ratio had fairly consistent effects on conflict: eight of the ten 80P data points here were higher than the corresponding 80CA data points. It therefore appears that a competitive orientation induced by gearing payments mainly to persuasion rather than to collective accuracy has a greater effect on conflict in CD than in CS pairs, especially at the outset of the judgment task.

Accuracy. Collective accuracy, measured in terms of the discrepancies between *S*s' joint judgments and the actual values of the

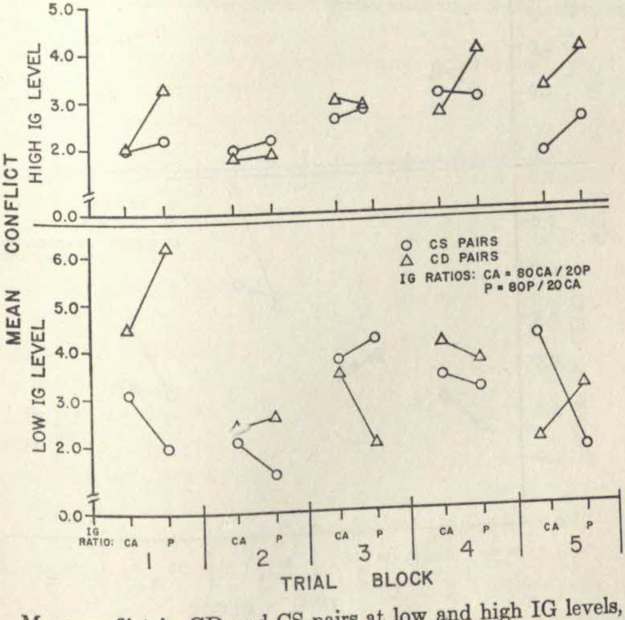


FIG. 2. Mean conflict in CD and CS pairs at low and high IG levels, plotted as a function of IG ratios within trial blocks. Each data point represents the mean of five pairs of subjects.

criterion, is shown in the form of error scores plotted in Fig. 3. The results of the four-way ANOVA on these scores indicate that of all possible main effects and their interactions, only two were significant at even a borderline level: the main effect of trial blocks [$(F = 37.75)$, $p < .001$] and the main effect of IG ratios [$(F = 7.01)$, $p < .05$]. The data plotted makes interpretation of the former obvious: errors were small at the outset, grew larger as *S*s sought perfect accuracy, and became small again as they learned the nature of the uncertain task. The more important result concerns IG ratios. Errors for all pairs were significantly greater in the 80P than in the 80CA condition. This result is reasonably clear in the data plot, but it does not seem dramatic because the variability of the error scores was very high. However, the major

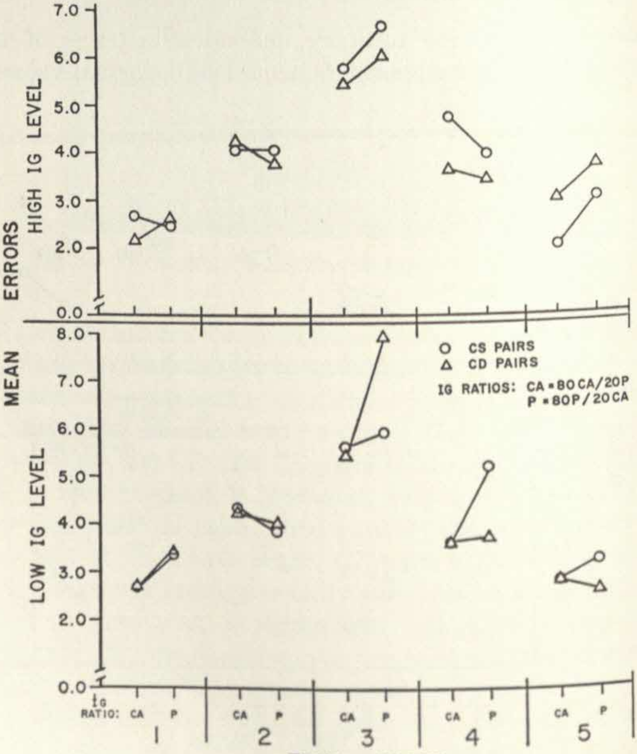


FIG. 3. Mean judgmental error scores in CD and CS pairs at low and high IG levels, plotted as a function of IG ratios within trial blocks

finding here is that the competitive orientation induced by the 80P condition impeded collective judgmental accuracy.

In general, our results show that when the cognitive-conflict model was modified to create a mixed-motive situation, the resulting conflict was most severe for persons who initially thought differently about their mutual problem and who were working under conditions where payoffs were high and geared to a competitive reward system. The latter factor also reduced collective accuracy. These findings are relevant to—but more complex and less pessimistic than—those developed in research based on the model for nonzero-sum games. The long runs of conflict typically found in gaming studies did not here occur. Instead, conflict appeared to vary as a complex function of reward structures, the discrepant meanings people attribute to cues, and the amount of experience people have on the experimental task.

NOTES AND DISCUSSION

GOLDEN SECTION: REASSESSMENT OF THE PERIMETRIC HYPOTHESIS

J. M. HINTZ and T. M. NELSON, *University of Alberta*

Abstract. By employing modern ophthalmological equipment, this study reassessed the hypothesis that the shape of the binocular visual field determines aesthetic preference for rectangular forms approximating golden-section dimensions. It also reexamined the finding that aesthetically preferred rectangles are oriented in correspondence with the horizontal orientation of the binocular visual field. The results indicate that the proportions of the binocular visual field should not be used to explain preference for rectangles having golden-section characteristics. They also question the relation thought to exist between preferred orientation of rectangles and orientation of visual field.

Stone and Collins have attempted to explain preference for rectangular forms approximating golden-section dimensions.[1] According to them, properly oriented rectangles with such proportions contain at least 90% of the area within the outline of the binocular visual field (the natural binocular field subtends 200° laterally and 130° vertically).[2] Preference presumably arises because such forms fit the natural binocular condition.

But when Schiffman studied the relation between the preferred orientation of drawn rectangles and the horizontal orientation of the binocular visual field, he had some difficulty eliciting drawings of rectangles approximating the golden section.[3] Preferred rectangular width-length ratios were less than the golden section, with a modal value between .459 and .560 and a median value of .500. Aside from this, however, he confirmed the predicted relation between response orientation and visual-field orientation. He also

---

Received for publication September 8, 1969.

[1] L. A. Stone and L. G. Collins, The golden section revisited: A perimetric explanation, this JOURNAL, 78, 1965, 503–506.

[2] D. O. Harrington, *The Visual Fields: A Textbook and Atlas of Clinical Perimetry,* 2nd ed., 1964.

[3] H. R. Schiffman, Golden section: Preferred figural orientation, *Percept. & Psychophys.,* 1, 1966, 193–194.

suggested reexamination of the perimetric hypothesis using other measures of the binocular field.

The present paper reports such a reassessment employing modern ophthalmological equipment. If the perimetric hypothesis is valid, individual preferences ought to shift in accord with the overall form of the functional binocular field.

METHOD

Subjects. Twenty sighted *S*s were drawn from University of Alberta students and University of Alberta nonacademic employees. The age range of the sample was 18 to 61 years, the mean age being 32.45. There were 10 men and 10 women. No visual pathology was reported that was not due to refractive media.

Apparatus. Rectangular outlines with constant areas of 81 sq. cm. and width-to-length ratios of .10, .20, .30, .40, .45, .50, .55, .60, .65, .70, .75, .80, .90, and 1.00 served as targets. The rectangles were drawn with black ink on 10 × 12 in. off-white paper. A Hans Goldman perimeter (manufactured by Haag-Streit Co., Berne, Switzerland) was employed to measure peripheral visual fields.[4]

Procedure. Targets were presented from a distance of 40 in. behind a 24 × 25 in., black field stop. Peepholes were cut into this so as to allow binocular viewing with eyes fixed at a height representing the center point of the target. The target was illuminated by fluorescent fixtures providing 23 ftc. on the target surface.

The *S*s were comfortably seated in front of the field stop and instructed. Briefly, they were told that the experiment was an investigation into preference for rectangles which would vary in proportion. They were also told they would be presented with three cards on a number of occasions and should choose the rectangle they found most aesthetically pleasing, that next most pleasing, and the least pleasing. The aesthetic experience was stressed and no time limits indicated.

Aesthetic preference for the chosen rectangles was determined by a process of successive approximations. The *S*s were initially presented with rectangles having ratios of .30, .60, .80, and their first, second, and third choices were recorded. The criteria for preferences in the first set was the rectangle chosen two-thirds of the time in a minimum of six trials. Presentation positions were randomized. Following this, a new set of rectangles representing proportions falling in the range of the first and second choices was offered, and the same choice procedure used. Using this technique, it was found possible to determine each *S*'s preference within a range of five rectangles by Trial Block 12. These five rectangles were then presented three at a time in all possible combinations for 10 trials. The rectangle preferred most frequently in the final trial block was taken to represent *S*'s preference. After preference was determined by this method, *S*s were asked to draw their most aesthetically pleasing rectangle. Drawings were made with a pencil on 8 × 10 in. sheets of white paper. Corrections were permitted, and again the aesthetic component of the response was stressed. Finally, each *S* was instructed to orient his drawn rectangle to a position considered most aesthetically pleasing.

---

[4] For a complete description, see Harrington (n. 2), p. 31.

Following experimentation, visual fields were measured by a professional clinical perimetrist.[5] The technique employed involved viewing a field in which a round white target ($\frac{1}{4}$ mm.^{2} in diameter and transmitting 100 mL.) moved in and out of the visual field. The *S*s were examined at 48 peripheral points, each separated by intervals of $7\frac{1}{2}°$, and the target was moved clockwise and from outside the binocular visual field into the binocular visual field. Each *S* reported when he first saw the object come into view.

RESULTS AND DISCUSSION

Chosen and drawn rectangles compared closely in value but ranged in width-length ratios from .20 to 1.00 and from .19 to .98 respectively. In general, the width-length ratios of chosen and drawn preferences were less than the golden-section ratio. The median for chosen preferences was .558 and the median for drawn preferences was .545. Since skewing toward lower values was present, the modal values tended to approximate golden-section dimensions more closely, particularly for the choice data. The mode was .60 for chosen preferences and .57 for drawn preferences. The mode included 28% and 20% of the cases respectively. In the limits .55–.65, which range includes the golden section, the percentage of *S*s became 40% in both cases.

Orientation data were not in close agreement with those of Schiffman and thus questioned the relation posited between response orientation and visual field orientation. Only 80% of the *S*s oriented their drawn rectangles horizontally, whereas Schiffman reported 97% correspondence. One *S* preferred a vertical orientation, and three *S*s oriented their drawings to the shape of a diamond.

Accepting the validity of these results, width-length ratios of the *S*s' chosen rectangles were correlated with the width-length ratios of their visual field's "average rectangle" (see n. 1). The coefficient of correlation calculated was .279, indicating little relationship between the variables. Further, inspection of the data shows that the ranges of the two variables were distinct. Preferences extended from .20 through to 1.00, whereas the average rectangle of the visual field was never less than .47 or greater than .84. When the data are plotted as in Fig. 1, the points fall in only a small portion of the graph rather than being distributed along the diagonal as predicted. Thus, the theoretical expectation was not born out. It does not appear likely that preferences for rectangles that approximate the golden section can be reduced to characteristics of the binocular visual field.

---

[5] The authors would like to thank Miss Barbara Salter, A.H.A., of the University of Alberta Hospital Glaucoma Clinic, for her cooperative spirit in measuring the visual fields.

In sum, aesthetically pleasing proportions of rectangles were ascertained using two procedures. Preferences obtained from one of these methods were then correlated with the average rectangle of the visual field, which was calculated from data provided by a professional clinical perimetrist. The measurements were unreliably correlated, which suggests that relation between the variables is spurious. It appears that the proportion of the binocular field should not be used to explain preferences for rectangles having golden-section characteristics. It also appears that the preferred horizontal orientation of rectangles is not as firm as earlier work suggests.

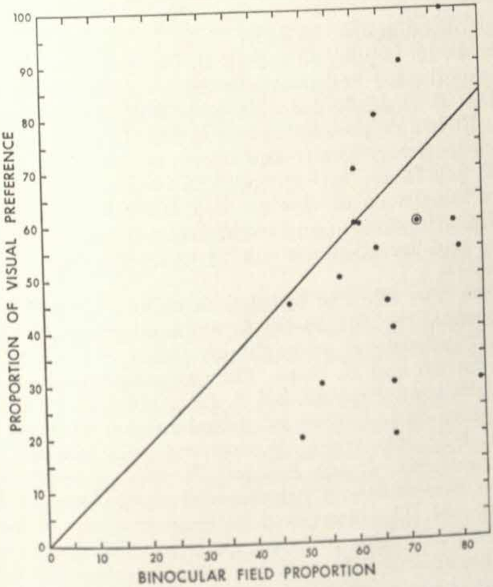


FIG. 1. Proportional preferences compared with proportions of binocular visual field. Perfect dependence of aesthetic preference upon the proportions of the binocular visual field is predicted by the solid diagonal line.

INSTRUCTIONS AND KNOWLEDGE OF THE SITUATION IN BRIGHTNESS PERCEPTION

MARTIN S. LINDAUER and ROGER F. BAUST,
State University College of New York at Brockport

Abstract. The influence of information and instructions on brightness perception was investigated under conditions where either ratio or retinal matches were possible. Unlike most earlier studies, this one found no effect of experience: both trained and control groups showed identical ratio responses. The importance of obtaining perceptual rather than cognitive reports is emphasized in accounting for the apparent contradictions in the literature.

Experience and training have been shown to influence the perception of shape, size, and brightness.[1] It is of particular note that constancy for each has also been shown to be affected by instructions requiring observers to make retinal rather than phenomenal reports of their observations.[2] The susceptibility of such basic phenomena to learning has been argued strenuously by the transactionalists and to a lesser degree, by gestaltists as well.[3] But it is

---

Received for publication July 22, 1969.

[1] On shape, see P. K. Leppmann and R. B. Mefferd, Jr., Validity of perceptual reports of experienced and inexperienced observers, *Percept. mot. Skills*, 26, 1968, 1167–1172; R. B. Mefferd, Jr., Betty A. Wieland, D. G. Greenstein, and P. K. Leppman, Effects of pretraining and instructions on validity of perceptual reports by inexperienced observers, *Percept. mot. Skills*, 27, 1968, 1003–1006. On size, see H. R. Schiffman, Size-estimation of familiar objects under informative and reduced conditions of viewing, this JOURNAL, 80, 1967, 229–235. On brightness, see R. G. Hopkinson, An experiment on the assessment of brightness under 'free-choice' and 'forced-choice' conditions by a group of observers, *Ergonomics*, 3, 1960, 44–50.

[2] A. A. Landauer, The effect of instructions on the judgment of brightness, *Quart. J. exp. Psychol.*, 16, 1964, 23–29; A. A. Landauer and R. S. Rodger, Effect of "apparent" instructions on brightness judgments, *J. exp. Psychol.*, 68, 1964, 80–84; M. Parrish and K. Smith, Simultaneous brightness contrast as a function of perceptual set, *Psychon. Sci.*, 7, 1967, 155–156; A. A. Landauer, Influence of instructions on judgments of unfamiliar shapes, *J. exp. Psychol.*, 79, 1969, 129–132; W. H. Lichte and C. R. Borresen, Influence of instructions on degree of shape constancy, *J. exp. Psychol.*, 74, 1967, 538–542; V. R. Carlson and E. P. Tassone, Size-constancy judgments at equal distances, *Percept. mot. Skills*, 27, 1968, 193–194. The terms used in the instructions of these studies to contrast the two attitudes have included: analytic, critical, objective, physical, projective, real, vs. apparent, immediate, naïve, perceived, synthetic, whole.

[3] On the transactionalists, see W. H. Ittelson, The constancies in perceptual theory, *Psychol. Rev.*, 58, 1951, 285–294; W. H. Ittelson and F. P. Kilpatrick, Experiments in perception, *Sci. Amer.*, 185, 1951, 50–55; F. P. Kilpatrick and W. H. Ittelson, The size-distance invariance hypothesis, *Psychol. Rev.*, 60, 1953,

surprising to find brightness included among those phenomena shown to be affected by experience. MacLeod's early investigation did not find it so, nor would such an effect seem consistent with the great body of work which places brightness perception within a psychophysical framework.[4] Furthermore, in contrast to shape and size, brightness is thought of as being more stimulus-bound and dependent upon retinal factors.[5] Thus, the treatment of brightness and brightness constancy as a function of ratios of intensities explicitly rules out the traditional explanation of 'taking illumination into account,' *i.e.* denies the role of knowledge.[6] Finally, everyday observations of objects in different illuminations (and the observations in the Gelb and Katz demonstrations) do not appear to be events easily amenable to phenomenal change as a result of knowledge of the situation or of conscious attempts to reverse one's percept.[7]

The purpose of this study was to investigate whether maximizing observers' critical understanding and approach to a brightness-perception task would indeed lead to a change in the outcome. To this end, two groups differing in their knowledge of the situation and in their instructions were asked to make brightness matches under ratio conditions of illumination. Thus, the structural factors thought responsible for brightness perception would be shown to persist if observers made ratio rather than retinal matches despite circumstances which favored the opposite result.

METHOD

As part of their research requirements during the beginning weeks of a perception course, 41 undergraduate *O*s were randomly divided into two groups,

---

223-231. On the gestaltists, see K. Koffka, *Principles of Gestalt Psychology*, 1935, 235; K. Duncker, The influence of past experience upon perceptual properties, this JOURNAL, 13, 1939, 255-265.

[4] R. B. MacLeod, Brightness constancy in unrecognized shadows, *J. exp. Psychol.*, 27, 1940, 1-22; R. B. Freeman, Jr., Contrast interpretation of brightness constancy, *Psychol. Bull.*, 67, 1967, 165-187; J. L. Brown and C. G. Mueller, Brightness discrimination and brightness contrast, in C. H. Graham (ed.), *Vision and Visual Perception*, 1966, 208-250.

[5] R. H. Forgus, *Perception*, 1966, 53-54, 75-76, 98-103; H. Wallach, The perception of neutral colors, *Sci. Amer.*, 208, 1963, 107-116.

[6] H. Wallach, Brightness constancy and the nature of achromatic colors, *J. exp. Psychol.*, 38, 1948, 310-324; Wallach (see n. 5 above), 109.

[7] For Gelb, see A. Gelb, Colour constancy, in W. D. Ellis (ed.), *A Source Book of Gestalt Psychology*, 1955, 196-209; also in Wallach (see n. 5), 108. For Katz, see Gelb, *ibid.*, 198-200; Koffka (n. 3), 241-242; Wallach (n. 5), 107.

trained ($N = 20$) and control ($N = 21$); there were 10 women in the former group and 14 in the latter. Although both groups were knowledgeable about the psychophysics and sensory aspects of perception, as well as shape perception, the topic of the constancies or of brightness had not yet been presented. Prior to the experiment, the trained (informed and instructed) group received a 20-min. lecture and discussion on brightness perception and constancy, including an explanation in terms of ratios of intensities and a demonstration of the apparatus to be used in the experiment.[8] In addition, at the time of the experiment, they were told to make brightness matches that would offset any ratio effect by concentrating on the inner fields of discs of the standard and variable panels, and try to ignore the surrounding outer fields or rings of light, *i.e.*, to make a retinal match. The control (uninformed and uninstructed) group was given no special preparation or instructions, but simply told that they were participating in an experiment in which they were to match the brightness of two discs until they appeared to be the same, *i.e.* to make a phenomenal match. (In previous demonstrations and experiments to which *O*s had been exposed, the importance of immediate and spontaneous reports had been emphasized.)

The panels containing the two sets of standard and variable rings and discs were slightly separated from each other; there was a distance of 26 in. between their centers. A black barrier was extended between the panels to ensure the separation of the electroluminescent-light sources, which could be independently and continuously varied. The discs were 4⅛ in. wide, the contiguous rings 1⅞ in. wide; their overall diameter was 7⅞ in. The study was conducted in a dark room, in two separate group sessions. There were four settings of the ring and disc of the standard (in ftL.): 2.0:1.0, 3.0:1.0, 4.0:1.0, and 5.0:1.0 (*i.e.* ratios of 2, 3, 4, and 5:1 respectively). Corresponding to each of these values, the ring of the variable was set at 4.0, 4.4, 1.6, and .5 respectively. Consequently, retinal matches in the variable disc would be values of 1.0 in all four cases, while ratio matches would be 2.0, 1.5, .4, and .1 respectively, *i.e.* equivalent to the ratios present in the standard. One ascending and one descending trial alternated for each of the four settings, randomly presented by the method of limits. Beginning well below or above the limits for both retinal and ratio settings of the variable disc, *E* manipulated and called out the dial values at intervals of approximately .2 ftL.; these units were arbitrarily related to the actual luminescences involved. In order to approximate the conditions usually found in individual testing, *O*s were told to note and remember that value which designated a brightness match, and to write it down on their data sheets only at the end of the trial. They were cautioned not to change the initial value chosen even if subsequent settings appeared to be better choices. Postexperimental interviews indicated that these procedures were followed. The room lights were turned on at the end of each trial. The distance between the nearest *O* and the panels was 6 ft.; for the *O* sitting farthest away, the distance was 18 ft. *O*s were no more than 3.5 ft. (for near *O*s) or 5 ft. (for far *O*s) away from a central view of the apparatus.

---

[8] The explanation was based on Wallach (nn. 5 and 6). The apparatus was Brightness Comparator V-0659, Polymetric Co., 1415 Park Avenue, Hoboken, N. J. 07030.

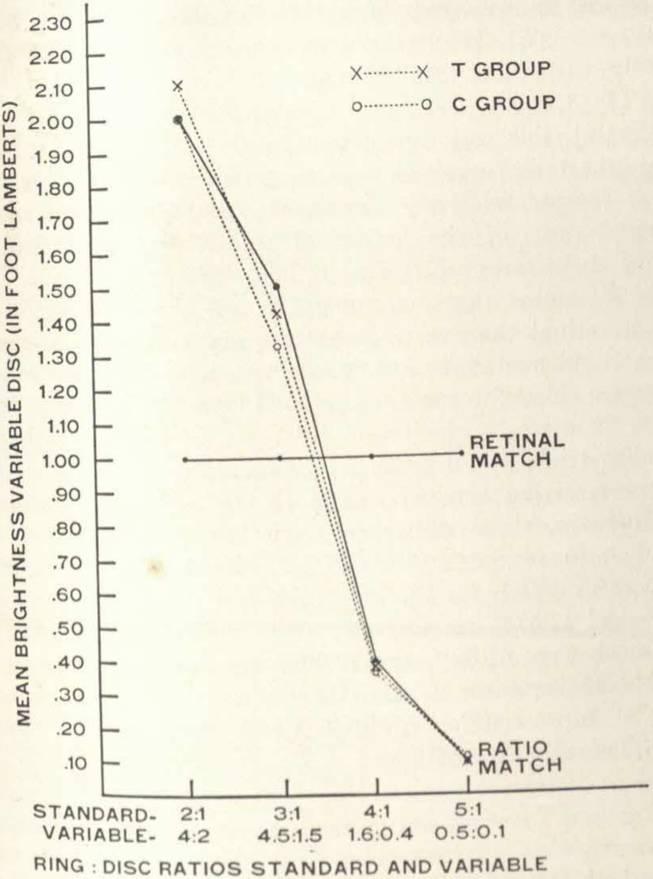


FIG. 1. Brightness settings of trained (T) and control (C) groups

RESULTS

As Fig. 1 indicates, both trained and control groups made brightness matches which were nearly identical with ratio expectations and quite divergent from retinal values. The overall mean (and *SD*) percentage deviations from a ratio match by the trained and control groups were 13.40% (8.17) and 12.43% (7.60) respectively, a difference which was not significant [t (39) = .38, $p > .05$]. (In relation to retinal predictions, the deviations of the trained and control groups were, of course, much larger: 76.37% vs. 72.87% respectively.) There was no difference between the two groups at each of the four settings, using the original ftL. data [t (39) ≤ 1.81, $p > .05$]. While the deviations from a ratio match significantly varied between settings for both trained and control groups [F (3, 57) $= 3.09$, $p < .05$ and F (3, 60) $= 6.50$, $p < .01$ respectively], this was contributed by the 4:1 and 5:1 positions. Because the units involved were so small (.1–.4 ftL.), small deviations became relatively large when converted to percentages. An examination of the individual scores confirmed that they clustered about ratio rather than retinal expectations. Thus, from a total of 164 scores, there were only 11 cases in which matches were closer to retinal than to ratio values: eight cases in the control group (all but one at the 3:1 setting, where variable and standard values were closest to one another) and three in the trained group. Finally, there was a significant difference between ascending and descending trials at all settings but one (5:1; $t < 1$), the latter sequence receiving brighter values [t (19 and 20) ≥ 2.56, $p < .05$]. However, these differences nonetheless remained closer to ratio than to retinal matches. The mean percentage deviations from a ratio match for the ascending and descending series were 17.83% and 11.67% respectively, while the deviations from a retinal match were 61.33% and 79.00% respectively. Further, there were but 15 instances (again, the majority of them at the 3:1 ratio), all for ascending trials, in which matches were closer to retinal than ratio expectations.

DISCUSSION

Since both trained and control groups fulfilled ratio rather than retinal predictions, it appears that informational indoctrination and instructional set to respond otherwise were ineffective in influencing brightness perception. The group and individual data were quite clear and consistent on this, even though but two measures were taken at each of the four settings and the less precise group procedures were used. Postexperimental inquiry also revealed that the trained observers struggled to make retinal matches, using such techniques as squinting and cupping their hands in order to make a reduction screen that would eliminate the surround.

In order to account for the results of the studies that did find instructions to have an effect on brightness perception, it should be noted that those studies did not inform observers about the *nature* of brightness perception, except to the extent of defining attitudes toward the task.[9] It is interesting that the only other study to show no effect of instructions was one which, like the present study, fully informed observers about the phenomenon.[10] Paradoxically, this implies that the perceptual set established by instructions alone must have been stronger than the set initiated by including information that would presumably lead to more complete comprehension. This anomalous situation suggests that the critical difference between positive and negative studies may be a matter of how the observers responded to the instructions. With little understanding of the phenomenon being observed, they may have come to rely on the experimenter's instructions and reacted in terms of what they thought was expected of them. They may, in effect, have thus produced cognitive rather than perceptual reports of brightness.[11]

While none of the studies in question provide information on the observers' qualitative reactions, Landauer appears aware of the issue.[12] That he found no significant intra-observer variability he took as indication of an absence of vacillation or uncertainty, which he then used as proof that the observers understood the task. Yet, according to the present argument, this stability of report may just as well have been due to consistency of judgment as to constancy of perception. The fact that studies have used a situation like Katz's (*i.e.* two compartments differing in illumination, see n. 7) lends itself to this argument. This is because the proximal stimulus is somewhat equivocal in the more realistic situation, since ratios of intensity are given indirectly, in terms of reflectance or albedo, rather than directly, as in the present instance. Consequently, the presence of a textured surface may yield small but varying intensities; and the spatial separation between interacting fields may reduce the necessity of their relationship. Because the surface colors are somewhat ambiguous, there is therefore more latitude for interpretation under conditions that experimentally induce judgments.[13]

---

[9] Those studies were the ones by Landauer, Landauer, and Parrish and Smith (see the first three references in n. 2).

[10] See MacLeod (n. 4 above).

[11] T. Natsoulas, Interpreting perceptual reports, *Psychol. Bull.*, 70, 1968, 575–591; T. Natsoulas, What are perceptual reports about? *Psychol. Bull.*, 67, 1967, 249–272; M. S. Lindauer, Reply to Harris' comments on "Experience with congruity in the perception of incongruity," *Psychon. Sci.*, 15, 1969, 79.

[12] Landauer (n. 2, second reference), 80–81.

[13] The authors wish to thank Dr. Wallach for suggesting some of these possibilities as contributing factors.

BOOK REVIEWS

Edited by JULIAN HOCHBERG
Columbia University

HUMAN INFORMATION PROCESSING: INDIVIDUALS AND GROUPS FUNCTIONING IN COMPLEX SITUATIONS. By H. M. SCHRODER, M. J. DRIVER, and SIEGFRIED STREUFERT. New York: Holt, Rinehart and Winston, 1967. Pp. viii, 224. $7.95.

This monograph describes a research program that emphasizes the abstract structural properties of both conceptual systems and environmental situations. Such a program can study the interaction of environmental complexity (or stress) and level of conceptual structuring in a wide variety of content areas ranging from group to individual functioning in educational, game-playing, and other complex settings. The authors achieve such generality by focusing on the way individuals process information in a given situation rather than on the content of this information. For a given individual in a given situation, his way of handling the environmental input can be indexed, according to the authors, on a scale that varies from very low integrative complexity (concrete) to very high integrative complexity (abstract). Using a variety of indirect measures, the authors present data, on both groups and individuals, that support their updated version of the Yerkes-Dodson law: as environmental complexity or stress increases, the individual or group conceptual structure becomes progressively more abstract up to a maximum and then declines as the environmental overload increases. They further demonstrate, as hypothesized, that the difference between individuals who are initially more conceptually complex and those who are conceptually simple shows up more in the middle range of environmental complexity than at the extremes.

Both the strength and weaknesses of the research reported in this book stem from the fact that the authors have not yet committed themselves either to a clear formal definition of integrative complexity or to a standard procedure for measuring it in a given situation. This is a strength because it has enabled them to test their model in a variety of complex situations in which it would be impossible to directly measure abstractness of conceptual structures. Instead, they have been free to adapt the measuring instrument and indexes to each particular context. Sometimes their indexes are derived from sorting procedures, from paragraph completions and other projective devices, or from similarity judgments of pairs. Integrative complexity is inferred from subjective ratings of written products, from multidimensional scaling of similarity spaces, from flexibility in changing impression formations, and so forth. By avoiding premature commitment to one type of index or one form of assessment device, the authors have left themselves room for further explication of their construct in terms of consistent behavior of a multiplicity of indexes across a variety of settings.

The authors seem well aware of the weaknesses that are also inherent in this approach. When a number of their separate indexes behave the same way in a given study, the temptation is to use this as further support for their construct. When the different measures behave somewhat differently within the same study, as happens in some of the ones reported in the book, the authors are able to suggest provocative reasons for the different levels at which the indexes peak, and so on. And, finally, it is easier to find some subset of measures that will follow the U-shaped hypothesis in each different situation provided no commitment is made in advance as to which particular combination of indicators best reflects level of integrative complexity. Despite these weaknesses, the overall picture that one gets from the varied studies covered by the monograph is quite encouraging.

Although the idea behind this program is exciting and although the authors are apparently well on their way to developing feasible methods to study the interaction of environmental pressures and individual conceptual structures, this reviewer found the book rather difficult to read. I cannot decide whether this reflects a fault of the exposition, an inevitable consequence of presenting thinking and data that are still in progress, or the fact that the conceptual underpinnings of their approach just happen to be difficult to communicate. At any rate, the first 54 pages present the working model in an abstract manner with no aid either from concrete referents or from formal definitions. The reader is not told if this dimension of integrative complexity is a logical construct, a generalization from empirical data, or a refinement and extension of such previous ideas about abstract conceptual structures as those of Lewin, Piaget, Zajonc, and others. Some concrete paradigms would help the reader to better grasp the somewhat vague ideas behind this way of ordering information-processing strategies. At the other end, some formal machinery for defining dimensions, values on dimensions, combinatory rules, similarity spaces, and other elements of the model (machinery such as that used by Hunt and Restle, for example) would help to avoid ambiguities at various points in the discussion. It is not always clear, for example, whether integrative complexity implies a reordering of items on a given dimension, complex but consistent rules for combining dimensions, applying more than one type of combinatory rule to the same set of dimensions, spanning more elements with a given set of dimensions, or all of these. An early definition of degree of integrative complexity, for example, emphasizes the number of different combinatory rules that the individual can apply to the same set of dimensions in dealing with a given set of objects. Yet, one of the favored measures of integrative complexity derives from using multidimensional scaling methods that presuppose that the subject uses *one* combinatory rule consistently (usually the Euclidean rule). Indeed, in the studies using indexes derived from multidimensional scaling, one might argue that the data are more consistent with the idea that integrative complexity implies the ability to use *one* combinatory rule consistently across an entire domain of objects. A truly abstract subject, according to strict interpretation of the authors' definition, should produce a multidimensional solution that is pure hash.

Indeed, it seems that only the concrete end of the abstract-concrete scale is clearly defined. The lowest level of integrative complexity is exemplified by the individual who uses a small number of dichotomous dimensions in a rigidly fixed way in dealing with items in a given environment. As the number of dimensions a person employs increases and as he makes finer discriminations on each dimension, the opportunity and probability of his becoming integratively complex increases. The authors do not want to identify their construct in terms of an increase in differentiation and discrimination. Instead, they argue that integrative complexity requires flexibility in combining and recombining the differentiated dimensions according to a number of different combinatory rules. The major weakness in their presentation, however, is that they have failed to demonstrate convincingly that they are dealing with something more than differentiation or discrimination. All of their studies using multidimensional scaling, for example, are consistent with the idea that dimensionality behaves according to the manner they have hypothesized for integrative complexity; at the same time, these measures based on multidimensional scaling can be interpreted as inconsistent with the idea that flexibility in using combinatory rules behaves according to hypothesis.

A careful reading of both the text and the appendixes suggest that the authors are aware of most of the problems and difficulties in the pursuit of their goals. They are to be congratulated for not letting these difficulties hinder them from plunging ahead despite what would appear to be formidable obstacles. The results reported seem to justify their point of view. Their approach illustrates the integrative complexity and flexibility that psychology needs if it is to generate concepts and laws that truly generalize beyond specific situations and tasks.

RAY HYMAN, *University of Oregon*

CONTEMPORARY THEORY AND RESEARCH IN VISUAL PERCEPTION. Edited by RALPH NORMAN HABER. New York: Holt, Rinehart and Winston, 1968. Pp. xi, 814. $12.50.

It is exciting to compare this new collection of research reprints with its ancestor of the 1950s, Beardslee and Wertheimer's *Readings in Perception*. The study of perception has prospered since the time it was an exotic transplant from abroad; nourished by new roots in information processing and neurophysiology, it aspires again to the central place in psychology. The new collection shares only two articles with the old, and those prophetic: Bruner's "On Perceptual Readiness" and George Miller's "The Magical Number Seven, Plus or Minus Two." Missing from both books, but prominent in its citation in the new, is Garner, Hake, and Eriksen's development of the technique of converging operations.

Since there are no less than 79 studies reprinted here (plus an extraordinary one published for the first time), none more than 15 years old, it would be a thankless task to discuss specific articles. For the areas covered, however, the prospective reader can be assured that many of the seminal articles are here. These areas, briefly, are (1) the sophisticated new techniques and concepts of perception measurement, which are turning hitherto subjective notions —such as that of image—into well-anchored empirical constructs, *e.g.* signal-detection theory, operant discrimination, computer-generated stimuli; (2) the neuropsychology of perception; (3) what might best be called visual cognition, attention, pattern recognition, and the transformation of perception into memory; and (4) the development of perception in the young, and its possible parallel, adult adaptation to transformed optical environments.

This is definitely a book about perception and not sensation, as classically conceived: there is little here about color, or about scaling, or even about the eye, but a lot about motion and form. Space perception and the constancies are curiously de-emphasized, although there has been no lack of good studies in these areas recently. Animal studies of perception are not well represented. Social perception is totally absent; the reason given, that it is really a matter of judgment rather than perception, is not convincing.

This book will serve as the text for top-notch advanced courses and as a convenient anthology for professionals; it is definitely not 'supplementary readings' for an introductory course. The studies are reprinted in almost all cases whole, without any unifying editorial material. In his concise introduction, Haber boldly and wisely discusses philosophical as well as methodological issues; wisely too, he steers clear of the professional philosophical articles on perception.

JOHN C. HAY, *University of Wisconsin—Milwaukee*

THE SCIENCES OF THE ARTIFICIAL. By HERBERT A. SIMON. Cambridge: The M.I.T. Press, 1969. Pp. x, 123. $5.95.

This pleasant little book comprises the Karl Taylor Compton Lectures given in the spring of 1968 at the Massachusetts Institute of Technology by Herbert A. Simon. The intent of the book is to explicate the nature of "artificial" phenomena. This word refers to a very broad and consequently vague concept. The metaphor that interested me was the characterization of an artifact as an interface between an "'inner' environment, the substance and organization of the artifact itself, and an 'outer' environment, the surroundings in which it operates" (p. 7). A consequence of this conception is the teleonomic character of artificial systems. In this image Simon proposes that the rules and principles of the artificial as viewed at the interface will give us a deeper insight into the object of inquiry itself and into its relation to its environment.

The first three chapters of the book develop this hypothesis at three levels; the first is with respect to an obvious artifact, the digital computer. In this section the author discusses simulation by digital computer and the character of computer science, namely, whether it is a formal or an empirical science. Simon concludes that the computer as an artificial system can be made to resemble in behavior other structures. Insofar as the organization of the parts of a structure is the interesting feature of it, we know as much as may be of interest to a practical man. As for the analysis of the computer as an object, Simon argues that one must experiment on it, that its nature is not completely explicated by purely abstract considerations. One may question whether the failure of designers to produce flawless computers constitutes evidence for this proposition.

The second chapter proposes that thought is an artificial system and, therefore, that psychology is a science of the artificial. This section of the book summarizes a variety of experiments on memory, learning, and concept formation. It concludes that thought may be viewed as an artifact and that the concept of the artificial therefore provides a reasonable scheme for understanding how people think. In the course of the discussion, Professor Simon makes some good suggestions about how to do psychological experiments, and he offers additional support for his view that the time required to remember a chunk of information is constant.

The third section may be of less interest to psychologists. It is concerned with an analysis of the science of design. It offers to engineers ways to solve problems in design. The author subsumes the task of designing anything under the rubric of problem solving. He then uses his theory of problem solving to form principles of procedure for when one is faced with a problem of design.

The book is capped with a reprint of "The Architecture of Complexity," first published in 1962. This article represents Professor Simon's view that the analysis of complexity depends upon a theory of hierarchic structures. This section complements the first three chapters by serving as an anticipatory summary of some of the topics that were discussed. It may be of no small importance to be shown how broad conceptual ideas become concretized and made operational by an active scientist attending to problems of paramount importance.

EUGENE GALANTER, *Columbia University*

TOPICS IN INTERMEDIATE STATISTICAL METHODS. Vol. 1. By T. A. BANCROFT. Ames: Iowa State University Press, 1968. Pp. xiii, 129. $5.90.

The author of this brief, well-produced book is a prominent statistician with an obvious flair for communicating with the practitioner. In this first of a proposed two-volume series, he has chosen to concentrate on the problem of unequal cell frequencies in analysis-of-variance designs. Chapter 1 deals in some detail, but certainly without any theoretical distraction, with two-way classifications having unequal cell frequencies. Chapter 2, seven pages in length, extends the discussion to approximate methods of handling this case. Chapters 3 and 4 parallel the first two chapters for the case of three-way classifications; and Chapter 5, four pages in length, touches on extensions to *n*-way designs. Chapter 6 handles covariance analysis and ends with a section on how to manage with unequal cell frequencies. Chapter 7 is a lucid discussion of orthogonal polynomials and their uses; and Chapter 8, the final one, gives a useful coverage of competing multiple-comparison procedures.

Throughout this little book, the style is fluid and easy to follow. Examples are well chosen (albeit mostly from the field of biometrics) and helpful, and many useful references are provided. In addition, the reader is given the benefit of the author's experience as he compares alternative strategies and gives us his evaluative opinions. This is a text that could prove very useful for a short graduate course in psychology, with or without the addition of more psychologically oriented examples. I expect it will find its way to the shelves of many serious students and teachers, both because it covers material that is relevant to them and because it is aimed so directly at practitioners.

VICTOR E. MCGEE, *Amos Tuck School of Business Administration*
Hanover, New Hampshire

LAW, LIBERTY, AND PSYCHIATRY. By THOMAS S. SZASZ. New York: Collier, 1963. Pp. xi, 281. $2.45.

This book consists of a revised and adapted collection of Dr. Szasz's articles on psychiatry and the law, most of which have appeared previously in various journals, books, and periodicals. Several of the articles emphasize the problems psychiatrists face on the issue of commitment on the grounds of insanity. A second major theme concerns the difficulties centering about the testimony of psychiatrists in criminal felony cases where a plea of 'not guilty by reason of insanity' or a claim of 'mitigation of criminal responsibility by reason of mental disease' is involved.

It is somewhat difficult for the reader to come to grips with this book because it is largely an adaptation of previous articles and so is not a single unified work. At times, it is difficult to know whether the author is enlarging on, shifting, or supporting some point made in a different article than the one being read at a given point in the book. Nevertheless, these articles raise many important and often related problems in a sophisticated and logical manner. The existence of this book saves the trouble of having to hunt through the literature for the original articles. And Dr. Szasz discusses several important questions: the relationship of psychiatry to social control; the nature of psychiatry as a science; the role of psychiatry in serving the law and of the law in serving psychiatry and society; and perhaps most important of all, the relation of psychiatry to civil liberties.

Many of Dr. Szasz's concerns are ethical (normative) in nature. Basically he feels that psychiatrists should not be placed in the position of having to commit individuals to mental institutions if the individuals do not want to be there. He is not at all certain that being considered 'mentally ill,' a concept he finds useless in any case, should be a cause for involuntary incarceration. Further, he does not believe that responsibility is a matter of the presence or absence of 'mental problems.' Szasz seems to be saying that all behavior has causes and that people may be considered either not responsible for anything they do or, with equal ease, responsible for anything they do. The particular nature of causal factors, whether they involve stress, schizoid tendencies, or an impoverished home, are irrelevant to the issue of responsibility. Rather, he argues, let the law find a person guilty or innocent of the charge at issue and leave psychiatry out of it. After the decision, however, psychiatrists should be consulted as to the best thing to do with a person *if* he is found guilty. If something other than vengeance is desired, then psychiatry may have a lot to offer in terms of changing the 'criminal' into a more acceptable person. Treatment (rehabilitation), not judgment about responsibility or guilt, he says, is the job of psychiatry.

Szasz is not at all certain that institutions for the 'mentally ill' are good for people, nor does he feel that the large mental hospital is the place for those in society whom others do not want around. He is concerned with finding better alternatives to help those individuals who cannot get along with themselves and with the world about them, because he is concerned about the civil liberties of all human beings, even those whose behavior is considered quite peculiar by others.

Szasz devotes Part One of his book to "Psychiatry as a Science," but instead of defining what distinguishes psychiatry as a scientific enterprise, he uses this section to attack the concept of 'mental illness' as a designation for social deviates and to point out the inadequacies of psychiatric classification. Only in the "Summary and Conclusions" section of his work does he finally define the science of psychiatry. He defines psychiatry, in essence, as the science which studies and tries to change human behavior. Such a definition may make many psychologists, economists, political scientists, sociologists, and biologists uncomfortable inasmuch as they also have a stake in studying certain aspects of human behavior. Furthermore, not only psychiatrists but parents, teachers, the clergy, social workers, and employers, as well as some psychologists, spend considerable time trying to alter human behavior. Once Szasz defines 'mental illness' as a myth and sees the medical aspects of psychiatry as relevant to only a quite narrow range of human problems, he clearly is left with a discipline that has no specific scientific jurisdiction.

It may well be that psychiatry is not a science but rather a branch of medicine which best treats those interpersonal problems that are caused by disorders of the patient's autonomic and central nervous system. The other aspects of human life with which psychiatry currently tries to cope, if Szasz is right, might best be left to nonmedically oriented disciplines. It is not at all clear if Szasz has considered this ultimate implication of some of his positions.

Robert Singer, *University of California, Riverside*

DIMENSIONS OF READING DIFFICULTIES. By A. T. Ravanette. New York: Pergamon Press, 1968. Pp. xiv, 102. Cloth, $3.25; paperback, $2.00.

This book is an elementary discussion of the contexts in which reading failure may occur. The work is largely descriptive and focuses on how one may view reading difficulties in a child—from psychological, sociological, physiological, or educational viewpoints. Specific case histories are brought up, usually to show how a particular child's problem was misinterpreted because of a failure to consider one or more dimensions of reading difficulty.

This book is the first in a new series of monographs. The aim of each volume in the series is to present a multidisciplinary consideration of a particular topic (such as reading). In this context, the book was disappointing: it indicates a bias toward a psychosocial view of reading problems and tends to neglect other interpretations (such as neurological ones). Even beyond this, the book fails to present more than a commonsense approach to any dimension of reading difficulty. While it is useful to enumerate and define the dimensions, it would perhaps have been even more useful to explain how they might be used (a la Descartes) to identify and classify a reading problem. As it stands, there has been no synthesis of diagnostic method, and each case is left in the trial-and-error realm of empiricism.

Arthur Chernoff and Joanna Williams, *University of Pennsylvania*

LEARNING DISORDERS. Vols. 1 and 2. Edited by Jerome Hellmuth. Seattle: Special Child Publications, 1965 and 1966. Pp. 406 and 422. Paperback, $4.85 each.

The current development of clinical practice by educators, physicians, psychologists, and others is treated in some 32 chapters or papers. The term learning disability is used to refer to discrepancies between a child's "achievement level and his functioning capacity based on his mental abilities" (Vol. 1, p. 172). The concern is to provide differential diagnosis and treatment of the abilities and disabilities underlying learning. Implicit in almost all the chapters is the assumption that these disorders are often specific to learning, perception, motor ability, emotional problems, or some other cause. Speculative neurological and genetic interpretations abound. The evidence is clinical and educational, and includes case studies. The volumes accurately reflect the state of the art of the diagnosis and treatment of learning disabilities. The overview is important, as major programs have been expanding in colleges of education across the country.

The message of these volumes is clear. Contact with experimental psychology is sparse. Reinforcement techniques are not included, nor has the area benefited from systematic research as psychologists know it, except from some peripheral contact with psycholinguists. The neglect of individual differences in learning could scarcely have a more graphic demonstration. The lamentable state of sophistication in an area heavily supported by federal research funds is clear. It will not improve until the psychological experimentalist and the psychometrician combine their efforts and make learning disorders a domain of rigorous empirical research. In the meantime the present volumes will interest readers of the JOURNAL largely to document what needs to be done.

KEITH G. SCOTT, *University of Illinois, Urbana-Champaign*

DEPTH-ELECTROGRAPHIC STIMULATION OF THE HUMAN BRAIN AND BEHAVIOR. By C. W. SEM-JACOBSEN. Springfield, Ill.: Charles C Thomas, 1968. Pp. xv, 222. $15.50.

For many years the author of this book has been involved in the treatment of mental disorders, and more recently Parkinson's disease, by depth-EEG surgery through implanted electrodes. In a patient undergoing such treatment it is important first to stimulate the brain site being considered for destruction before an actual lesion is produced. The response to stimulation may indicate what the effect of the destruction will be and can thus give promise of a therapeutic outcome or warn the surgeon that important normal functions would be endangered. The bulk of the present volume is a summary of the results of such electrical stimulation through 2,659 electrodes implanted in 82 patients seen over an 11-year period.

In his introductory chapters, the author stresses the importance of maintaining procedures which will safeguard the rights of the patient and his family (*a*) to know the potentialities of the operation, good and bad, and (*b*) not to be subjected to studies which are unpleasant themselves or of no service to the goal of deciding upon the best treatment. The methods of procedure—in electrode implantation and in the study of stimulation—are outlined in some detail. There is a brief discussion of self-stimulation, with some intriguing comments on the usefulness of rate as a measure, the effects of double stimulation, and so forth, but no real data on this phenomenon are presented.

The effects of stimulation were coded into one or more of 99 different categories of response, which were considered to fall into 17 groups. Much of the volume consists of a survey of each of the 17 groups, with a discussion of the nature of the responses, the localization and nature of the stimuli which produce them, and the extent of correlation of other response groups with the one in question. There is an inherent limitation in the data due to the fact that electrodes were implanted only where it was hoped they would be therapeutically helpful; we can but praise the author for this. Nevertheless, some suggestive results are available.

There are very few surprises in the localization data unless one is looking for very strict localization of function or none at all. Thus, the appearance of colored balls, either moving or fixed, was elicited only from occipital stimulation, but a 'vibrating blinking' in front of the eyes was produced by points in the frontal region and near the third ventricle as well. One rather interesting finding is that changes in speech, including facilitation as well as arrest, were obtained from stimulation of the right hemisphere as well as the left.

The correlations (over electrodes) of responses of the various groups and categories provide some interesting ideas as well, but use of the material is rendered difficult because of some unfortunate features of the statistical treatment and presentation. Phi coefficients were computed between each pair of categories but they are not reported. Instead, the relationships are presented in terms of significance level. Since all comparisons had a common total n, the relative significance of various relationships does parallel their relative strengths. However, with the large n used (apparently 1,594, the number of electrodes giving any response), high significance could be achieved with very low values of ϕ. Not only are we invited to view with interest a relationship with a value as low as .064 (since this value gives a p of .01), but since the strongest significance level used was .0001, we are led to overlook any differences in strength of relationship in the wide range between .098 (which gives the p of .0001) and the top of the scale.

Potentially even more serious is the fact that apparently no correction was made for lack of continuity in these data. We are told, for example, that the relationship between aphasia and flushing has a p of .0001. However, only one (of the 1,594 effective electrodes) produced aphasia. Obviously the relationship suggested by the fact that this electrode was one of the 99 that elicited flushing must be assessed as having a significance of only about $\frac{1}{16}$, since there would be that probability by chance alone that the 'aphasia electrode' would be among the 'flushing electrodes.'

In point of fact, two redeeming aspects must be stressed. (1) The summary data are presented in sufficient detail that an interested reader can recalculate the correlations and their significance. (2) The principal conclusions concerning the relationship between responses are given in terms of the major groups of response: since these groups were elicited by reasonably large numbers of electrodes, the failure to correct for noncontinuity is of little importance. What then are these relationships? Motor responses, changes in tremor, modifications of rigidity, and peripheral sensory changes are generally highly intercorrelated. Another cluster of groups finds cardiovascular changes, vegetative responses, and changes in consciousness closely related to one another. Changes in speech form something of a bridge, being highly correlated with groups in this latter cluster and moderately correlated with changes in rigidity. Mood changes have a different pattern; they have weak positive relationships with the second cluster of groups and weak negative correlations with the motor (etc.) cluster. These findings would be relatively trivial if they reflected solely gross anatomical localizations; *e.g.* the fact that many of the responses in the different groups of the motor cluster came from stimulation in or near the internal capsule (as they did), the fact that most of the autonomic-consciousness responses followed frontal stimulation (as they did), and so forth. However, it is reported that the same interrelationships described above were noted within each of four main sections of the brain when they were analyzed separately.

The data do indeed, as in the author's words, provide knowledge about "the functioning of the whole brain as a network of closely interrelated and interacting systems." This book is not one that everyone interested in neuropsychology must own. Nonetheless, it is well worth a casual reading to note the variety of responses produced and consider their relationships.

William A. Wilson, Jr., *University of Connecticut*

THE SKIN SENSES: PROCEEDINGS OF THE FIRST INTERNATIONAL SYMPOSIUM ON THE SKIN SENSES. Edited by D. R. Kenshalo. Springfield, Ill.: Charles C Thomas, 1968. Pp. xvii, 636. $26.00.

This symposium is a sample of current research on the skin senses and, as such, covers a range of topics from anatomy and physiology of receptors, through central processes, to psychophysical and perceptual problems. Some indication that the research is current may be gained by noting that in the proceedings of this symposium, held in June of 1966, over half of the more than 800 references are dated 1960 or later.

What advantages were there in printing this symposium rather than simply having the authors report their research in journals? First, in addition to presenting their current research, the authors here summarize their work for the past seveal years. Second, the authors were here permitted to discuss the significance of their results and how these results bear on some of the larger issues in the field, *e.g.* specificity vs. nonspecificity of receptors. Most of the authors made good use of the opportunity. Third, such a symposium reveals areas about which little is known and provides the impetus for future investigations. This end is achieved in part by the excellent discussions following each paper. One such area to which the discussants returned several times is the physical nature of the skin itself. What are the kinds of energy transformations that the skin imposes upon an electrical, mechanical, or thermal stimulus before such a stimulus even reaches a receptor? The answer to this question would be an important step in understanding such problems as the nature of the adequate stimulus for the different skin modalities and the correlation of psychophysical and electrophysiological data. The manner in which the authors made use of these advantages and the overall quality of the work reported ought to leave readers awaiting with interest the appearance of the *Second International Symposium on the Skin Senses.*

James C. Craig, *Indiana University*

BOOKS RECEIVED

(The books listed here have not as yet been noted in our pages. Listing here does not, however, preclude their later review.)

Agnew, N., and Pyke, S. *The Science Game: An Introduction to Research in the Behavioral Sciences*. Englewood Cliffs, N. J.: Prentice-Hall, 1969. Pp. 118. $3.50.

Alexander, Theron. *Children and Adolescents*. New York: Atherton Press, 1969. Pp. 365. $8.95.

Arnoff, F. N., Rubinstein, E. A., and Speisman, J. C., eds. *Manpower for Mental Health*. Chicago: Aldine, 1969. Pp. 204. $6.95.

Arnold, William J., ed. *Nebraska Symposium on Motivation, 1968*. Lincoln: University of Nebraska Press, 1968. Pp. 337. Cloth, $6.95; paperback, $3.25.

Ausubel, David P. *Readings in School Learning*. New York: Holt, Rinehart and Winston, 1969. Pp. 513. $4.95.

Ausubel, David P., and Robinson, F. G. *School Learning: An Introduction to Educational Psychology*. New York: Holt, Rinehart and Winston, 1969. Pp. 691.

Bandura, Albert. *Principles of Behavior Modification*. New York: Holt, Rinehart and Winston, 1969. Pp. 677. $9.95.

Bartley, S. Howard. *Principles of Perception*. 2nd ed. New York: Harper & Row, 1969. Pp. 578.

Beech, H. R. *Changing Man's Behavior*. Baltimore: Penguin Books, 1969. Pp. 283. Paperback, $1.95.

Bellak, Leopold, and Barten, Harvey H., eds. *Progress in Community Mental Health*. New York: Grune & Stratton, 1969. Pp. 271.

Birnbaum, N. *The Crisis of Industrial Society*. New York: Oxford University Press, 1969. Pp. 185. Cloth, $4.95; paperback, $1.75.

Bischof, Ledford J. *Adult Psychology*. New York: Harper & Row, 1969. Pp. 310. $5.00.

Bliss, Eugene L., ed. *Roots of Behavior: Genetics, Instinct, and Socialization in Animal Behavior*. New York: Hafner, 1969. Pp. 339. $15.00.

Borgatta, E. F., ed. *Social Psychology: Readings and Perspective*. Chicago: Rand McNally, 1969. Pp. 740.

Braginsky, Benjamin M., and others. *Methods of Madness: The Mental Hospital as a Last Resort*. New York: Holt, Rinehart and Winston, 1969. Pp. 225. $3.95.

Breger, Louis, ed. *Clinical-Cognitive Psychology*. Englewood Cliffs, N. J.: Prentice-Hall, 1969. Pp. 229. $7.95.

Brockman, J. *By the Late John Brockman*. New York: Macmillan, 1969. Pp. 166. $6.95.

Broderick, C. B., and Bernard, J., eds. *The Individual, Sex, and Society*. Baltimore: The Johns Hopkins Press, 1969. Pp. 406. $4.50.

Button, Alan DeWitt. *The Authentic Child*. New York: Random House, 1969. Pp. 237. $5.95.

Chauncey, H., ed. *Soviet Preschool Education*. Vol. I, *Program of Instruction*. New York: Holt, Rinehart and Winston, 1969. Pp. 182. $2.95.

Chertok, Leon. *Motherhood and Personality*. Philadelphia: Lippincott, 1969. Pp. 303. $10.50.

Cohen, Guido B. *The Task-Tuned Organization of Groups*. Amsterdam: Swets & Zeitlinger N. V., 1969. Pp. 212.

Cronbach, Lee J., and Gleser, Golding C. *Psychological Tests and Personnel Decisions*. 2nd ed. Urbana: University of Illinois Press, 1969. Pp. 347. Cloth, $7.95; paperback, $2.95.

D'Amato, Gabriel. *Residential Treatment for Child Mental Health*. Springfield, Ill.: Charles C Thomas, 1969. Pp. 186.

Dohrenwend, Bruce P., and Dohrenwend, Barbara Snell. *Social Status and Psychological Disorder: A Causal Inquiry*. New York: Wiley, Interscience, 1969. Pp. 207. $9.50.

Duncan, Hugh D. *Symbols and Social Theory*. New York: Oxford University Press, 1969. Pp. 314. $7.95.

Edel, Leon. *Henry James*. Vol. 4, *The Treacherous Years, 1895–1901*. Philadelphia: Lippincott, 1969. Pp. 381. $10.00.

Eron, L. D., and Callahan, R., eds. *The Relation of Theory to Practice in Psychotherapy*. Chicago: Aldine, 1969. Pp. 252. $5.95.

Evans, W. O., and Kline, N. S. *The Psychopharmacology of the Normal Human*. Springfield, Ill.: Charles C Thomas, 1969. Pp. 252. $11.50.

Fair, Charles M. *The Dying Self*. Middletown, Conn.: Wesleyan University Press, 1969. Pp. 240. $7.95.

Fairweather, George W., and others. *Community Life for the Mentally Ill*. Chicago: Aldine, 1969. Pp. 357. $10.00.

Ford, Laura May. *Bittersweet: The Autobiography of a Lesbian*. New York: Exposition Press, 1969. Pp. 144.

Fox, David J. *The Research Process in Education*. New York: Holt, Rinehart and Winston, 1969. Pp. 758. $8.95.

Gale, R. F. *Developmental Behavior*. New York: The Free Press, 1969. Pp. 600. $8.95.

Gamson, W. A. *Simsoc: Simulated Society: Participant's Manual*. New York: The Free Press, 1969. Pp. 108. Paperback, $3.95.

Garattini, S., and Sigg, E. B., eds. *Aggressive Behavior*. New York: Wiley, Interscience, 1969. Pp. 387.

Gazda, George M., ed. *Theories and Methods of Group Counseling in the Schools*. Springfield, Ill.: Charles C Thomas, 1969. Pp. 220. $7.50.

Goff, Beth. *Where is Daddy?* Boston: Beacon Press, 1969. Pp. 28. $3.95.

Gross, Charles G., and Zeigler, H. Philip, eds. *Readings in Physiological Psychology: Learning and Memory*. New York: Harper & Row, 1969. Pp. 313. Paperback, $5.00.

Gross, Charles C., and Zeigler, H. Philip, eds. *Readings in Physiological Psychology: Neurophysiology and Sensory Processes*. New York: Harper & Row, 1969. Pp. 286. Paperback, $5.00.

Haber, R. N., ed. *Information-Processing Approaches to Visual Perception*. New York: Holt, Rinehart and Winston, 1969. Pp. 418. $10.95.

Harshman, Hardwick W. *Educating the Emotionally Disturbed*. New York: Crowell, 1969. Pp. 490. Paperback, $5.50.

Hokanson, J. E. *The Physiological Bases of Motivation*. New York: Wiley, 1969. Pp. 182. Cloth, $8.95; paperback, $4.95.

Holubar, Josef. *The Sense of Time*. Cambridge: The M.I.T. Press, 1969. Pp. 122. $5.95.

Hooker, Davenport. *Prenatal Origin of Behavior.* New York: Hafner, 1969. Pp. 143. $6.00.

Horrocks, John E. *The Psychology of Adolescence.* 3rd ed. Boston: Houghton Mifflin, 1969. Pp. 628. $9.50.

Joffe, J. M. *Prenatal Determinants of Behavior.* Oxford: Pergamon Press, 1969. Pp. 366. $13.00.

Johnson, M. S., and Kress, R. A., eds. *Developmental Reading: Diagnostic Teaching.* Philadelphia: Temple University Press, 1968. Pp. 140.

Jones, Kenneth L., Shainberg, Louis W., and Byer, Curtis O. *Drugs and Alcohol.* New York: Harper & Row, 1969. Pp. 132. $2.50.

Jones, Kenneth L., Shainberg, Louis, and Byer, Curtis O. *Sex.* New York: Harper & Row, 1969. Pp. 168. $2.75.

Kantor, J. R. *The Scientific Evolution of Psychology.* Vol. 2. Chicago: Principia Press, 1969. Pp. 427. $12.00.

Keller, Fred S. *Learning: Reinforcement Theory.* 2nd ed. New York: Random House, 1969. Pp. 82. $1.50.

Köhler, Wolfgang. *The Task of Gestalt Psychology.* Princeton, N. J.: Princeton University Press, 1969. Pp. 164. $6.50.

Kutscher, A. H., ed. *Death and Bereavement.* Springfield, Ill.: Charles C Thomas, 1969. Pp. 364.

Lantz, H. R., and Snyder, E. D. *Marriage.* New York: Wiley, 1969. Pp. 512. $8.95.

Lehram, Daniel S. *Advances in the Study of Behavior.* Vol. 2. New York: Academic Press, 1969. Pp. 396, $15.00.

Levine, Jacob, ed. *Motivation in Humor.* New York: Atherton Press, 1969. Pp. 182. Cloth, $6.95; paperback, $2.95.

Levy, J. C. *Le temps psychologique.* Paris: Dunod, 1969. Pp. 106.

Lofland, John. *Deviance and Identity.* Englewood Cliffs, N. J.: Prentice-Hall, 1969. Pp. 330. Paperback, $4.95.

Luft, Joseph. *Of Human Interaction.* Palo Alto, Calif.: National Press, 1969. Pp. 177. $4.95.

Lundin, Robert W. *Personality: A Behavioral Analysis.* London: Macmillan, 1969. Pp. 463. $7.95.

MacAndrew, Craig, and Edgertown, Robert B. *Drunken Comportment: A Social Explanation.* Chicago: Aldine, 1969. Pp. 197. $5.95.

McNemar, Q. *Psychological Statistics.* 4th ed. New York: Wiley, 1969. Pp. 529. $10.95.

Maier, Henry W. *Three Theories of Child Development.* 2nd ed. New York: Harper & Row, 1969. Pp. 342. $8.25.

Martin, D., and Levey, A. B. *The Genesis of the Classical Conditioning Response.* New York: Pergamon Press, 1969. Pp. 145.

Marx, M. H. *Learning: Processes.* New York: Macmillan, 1969. Pp. 573. $9.95.

Medinnus, Gene R., and Johnson, Ronald C. *Child and Adolescent Psychology.* New York: Wiley, 1969. Pp. 692. $9.95.

Meetham, A. R., ed. in chief. *Encyclopedia of Linguistics, Information and Control.* Oxford: Pergamon Press, 1969. Pp. 718. $40.00.

Menninger, Karl. *Number Words and Number Symbols: A Cultural History of Numbers.* Cambridge, Mass.: The M.I.T. Press, 1969. Pp. 480. $15.00.

Michaux, William W., and others. *The First Year Out.* Baltimore: The Johns Hopkins Press, 1969. Pp. 298. $8.00.

Mills, Eugene S. *George Trumbull Ladd: Pioneer American Psychologist.* Cleveland: Case Western Reserve University Press, 1969. Pp. 299. $8.95.

Mills, Judson. *Experimental Social Psychology.* New York: Macmillan, 1969. Pp. 466. $8.95.

Miner, John B. *Personnel Psychology.* London: Macmillan, 1969. Pp. 310.

Morlan, George K. *Guide for Young Lovers.* Cranbury, N. J.: A. S. Barnes, 1969. Pp. 148. $5.95.

Munro, Alistair, and McCulloch, Wallace. *Psychiatry for Social Workers.* Oxford: Pergamon Press, 1969. Pp. 283. Cloth, $6.50; paperback, $4.75.

National Academy of Education. (Cronbach, Lee J., and Suppes, Patrick, eds.) *Research for Tomorrow's Schools: Disciplined Inquiry for Education.* London: Macmillan, 1969. Pp. 281. $6.95.

Oetting, E. R., and Thornton, George C., III. *Exercises in Psychological Testing.* New York: Harper & Row, 1968. Pp. 227. Paperback, $3.95.

Parsons, A. *Belief, Magic, and Anomie.* New York: Macmillan, 1969. Pp. 374. $7.95.

Pattison, E. M., ed. *Clinical Psychiatry and Religion.* Boston: Little, Brown, 1969. Pp. 327.

Perkell, Joseph. *Physiology of Speech Production.* Cambridge, Mass.: The M.I.T. Press, 1969. Pp. 104. $6.95.

Perls, Frederick, and Perls, S. *Gestalt Therapy Verbatim.* Lafayette, Calif.: Real People Press, 1969. Pp. 279. $5.00.

Plog, Stanley C., and Edgerton, Robert B., eds. *Changing Perspectives in Mental Illness.* New York: Holt, Rinehart and Winston, 1969. Pp. 757. $15.00.

Reid, Thomas. *Essays on the Active Powers of the Human Mind.* Cambridge, Mass.: The M.I.T. Press, 1969. Pp. 481. Paperback, $3.95.

Reid, Thomas. *Essays on the Intellectual Powers of Man.* Cambridge, Mass.: The M.I.T. Press, 1969. Pp. 481. Paperback, $2.95.

Roazen, Paul. *Brother Animal.* New York: Knopf, 1969. Pp. 211. $5.95.

Rollin, A. Robert, Smith, Jerome, and Wilson, William A., Jr. *Workbook and Laboratory Manual.* New York: Knopf, 1969. Pp. 306. $3.95.

Rollin, Henry R. *The Mentally Abnormal Offender and the Law.* Oxford: Pergamon Press, 1969. Pp. 139. Cloth, $4.75; paperback, $3.00.

Roscoe, John T. *Fundamental Research Statistics for the Behavioral Sciences.* New York: Holt, Rinehart and Winston, 1969. Pp. 336. $9.95.

Sadler, William A. *Existence and Love.* New York: Scribner's, 1969. Pp. 427. $10.00.

Salk, Lee, and Kramer, Rita. *How to Raise a Human Being.* New York: Random House, 1969. Pp. 205. $5.95.

Schmeiller, Gertrude, ed. *Extrasensory Perception.* New York: Atherton Press, 1969. Pp. 166. $2.95.

Schmitt, Francis O., Menlechuk, T., Quarton, Gardner, and Adelman, George, eds. *Neurosciences Research Symposium Summaries.* Vol. 3. Cambridge, Mass.: The M.I.T. Press, 1969. Pp. 594. $12.50.

Schulte, W., and Mende, W., eds. *Melancholie.* Herdweg, Germany: Georg Thieme Verlag, 1969. Pp. 236. DM 29,70.

Schutz, William C. *Joy: Expanding Human Awareness.* New York: Grove Press, 1969. Pp. 223. $1.25.

Scientific American. *Readings in Psychology*. Vols. 1 and 2. San Francisco: W. H. Freeman, 1969. Pp. 829. $20.00.

Shakow, David. *Clinical Psychology as Science and Profession: A Forty-Year Odyssey*. Chicago: Aldine, 1969. Pp. 350. $12.50.

Sherif, Carolyn, and Sherif, Muzafer. *Social Psychology*. New York: Harper & Row, 1969. Pp. 616. $11.95.

Sherif, Muzafer, and Sherif, Carolyn, eds. *Interdisciplinary Relationships in the Social Sciences*. Chicago: Aldine, 1969. Pp. 360.

Smith, Brewster M. *Social Psychology and Human Values*. Chicago: Aldine, 1969. Pp. 438. $12.50.

Smith, Charles P., ed. *Achievement-Related Motives in Children*. New York: Russell Sage Foundation, 1969. Pp. 263. $8.75.

Straus, Murray A. *Family Measurement Techniques*. Minneapolis: University of Minnesota Press, 1969. Pp. 316. $13.50.

Stulman, Julius. *Fields within Fields . . . within Fields: The Methodology of the Creative Process*. New York: World Institute Council, 1968. Pp. 122.

Swedenborg, Emanuel. *Angelic Wisdom concerning the Divine Love and the Divine Wisdom*. New York: Swedenborg Foundation, 1969. Pp. 306. Paperback, $1.00.

Tarnopol, Lester, ed. *Learning Disabilities*. Springfield, Ill.: Charles C Thomas, 1969. Pp. 389. $16.75.

Toman, W. *Motivation, Personlichkeit, Umwelt*. Göttingen: Verlag für Psychologie, 1968. Pp. 323. DM 28.

Trieschman, Albert E., and others. *The Other Twenty-three Hours*. Chicago: Aldine, 1969. Pp. 240. $7.00.

Watts, A. W. *The Two Hands of God*. New York: Collier, 1969. Pp. 236. Paperback, $1.50.

Weiss, Paul. *Sport: A Philosophic Inquiry*. Carbondale: Southern Illinois University Press, 1969. Pp. 274. $7.50.

Winick, Charles. *The New People*. New York: Pegasus, 1968. Pp. 384. $7.50.

Yablonsky, Lewis. *The Hippie Trip*. New York: Pegasus, 1968. Pp. 336. $6.95.

File No. 'Librar
Bureau Educ. Rev. Researc

PERCEPTUAL ORGANIZATION OF TWO-ELEMENT TEMPORAL PATTERNS AS A FUNCTION OF THEIR COMPONENT ONE-ELEMENT PATTERNS

David Preusser, W. R. Garner, and Richard L. Gottwald,
Yale University

Abstract. Subjects observed and then described continuously repeated two-element temporal patterns (one of two alternative elements per event) and their component one-element patterns (one element or its nonoccurrence per event). Descriptions could begin at any event in a pattern, since there are as many possible descriptions as events in continuous patterns. Descriptions of one-element patterns revealed two organizing principles: preference for longest gap at the end, or longest run of positive events at the beginning. Variability of pattern organization was low when in a particular pattern these two principles were compatible or when the principles were incompatible but with the gap principle stronger, and increased when the two principles were incompatible and equally strong or the run principle was stronger. The descriptions of two-element patterns showed that one element emerged as 'figure' and the organizing principles for that element determined the perceived organization for the two-element pattern. It was concluded that variability and difficulty of two-element perceptual organization is a function both of figure-ground relations between the two elements and of compatibility of one-element organizing principles.

A temporal pattern is a continuously repeating sequence of events. In a two-element pattern a single event is the occurrence of one of two alternative elements in a given unit of time. In a one-element pattern, an event consists of the occurrence or nonoccurrence of a single element. A two-element pattern may be thought of as the interlocking of two one-element patterns, since any two-element pattern can be decomposed into two unique one-element patterns. To illustrate, the two-element pattern XXXOXOO . . . contains XXX'X'' . . . and '''O'OO . . . as its unique one-element patterns, with the (') indicating the nonoccurrence of an element in the given unit of time.

There are as many different possible organizations for a temporal pattern as there are events in the pattern. For example, the two-element pattern XXO . . . can be organized XXO, XOX, or OXX. Each of these organizations, when repeated continuously, produces patterns which are indistinguishable from each other except in terms

---

Received for publication November 17, 1969. This study was supported by Grant MH-14229 from the National Institute of Mental Health to Yale University. The authors are indebted to Mr. Nigel Cox for help in instrumentation.

of the starting point for the sequence. Similarly, the one-element pattern XX′ . . . can be organized XX′, X′X, or ′XX. Previous studies on two-element patterns, however, have shown that subjects rarely use an organization which breaks a run of identical events,[1] so we can consider these runs as the meaningful psychological units. Patterns can therefore be coded in terms of sequences of run lengths. For example, the pattern XXXXOXOOO . . . can be represented as 4113, with alternative organizations of 1134, 1341, and 3411.

Previous work on temporal patterns has been concerned with three variables: pattern difficulty, perceived organization, and variability of organization (see n. 1). Our primary interest in the present paper is the perceived organization of temporal patterns; our secondary interest, variability of organization. Both the particular perceptual organization and the variability of organization should be based on a common set of principles, and discovery of these principles was the goal of the research reported here.

The first study dealing with perceived organization of these patterns was that of Royer and Garner.[2] They showed that when subjects have some degree of choice about which alternative organization they use, they do not use all possible organizations equally often, even when the organizations are specified solely in terms of run sequences. The patterns were eight events long, played auditorily at a rate of two events per second, and their subjects were required to follow or track the pattern on telegraph keys, with the tracking to begin after the patterns had been perceived with a definite organization. Their results showed that the subjects began to track (*a*) at the beginning of a run of three or more events, or (*b*) at the beginning of at least two cycles of a single alternation. Garner and Gottwald obtained similar results using verbal descriptions of auditory or visual patterns at various rates; so did Handel et al., using primarily measures of difficulty with mixed-modality presentations.[3] Thus, there are preferences for particular perceptual organizations—preferences which are quite general with respect to pattern presentation and type of response.

---

[1] See, for example, F. L. Royer and W. R. Garner, Perceptual organization of nine-element auditory temporal patterns, *Percept. & Psychophys.*, 7(2), 1970, 115–120.

[2] F. L. Royer and W. R. Garner, Response uncertainty and perceptual difficulty of auditory temporal patterns, *Percept. & Psychophys.*, 1, 1966, 41-47.

[3] W. R. Garner and R. L. Gottwald, The perception and learning of temporal patterns, *Quart. J. exp. Psychol.*, 20, 1968, 97–109; S. Handel and L. Buffardi, Pattern perception: Integrating information presented in two modalities, *Science*, 162, 1968, 1026–1028; S. Handel and W. E. Lewis, The

In a further study Royer and Garner used patterns of nine events so as to have a greater variety of possible organizations (see n. 1). Their results suggested that the preferred organizations are those which (*a*) are balanced in time with long runs at the ends of the pattern or (*b*) have a directional simplicity with run lengths either increasing or decreasing in regular order. These principles are generally consistent with, and represent an extension of, the results obtained with eight-event patterns. In general, Royer and and Garner considered these organizing principles as wholistic in nature, the patterns being perceived as totalities, with all parts interacting in a forward and backward fashion. The exact nature of this interaction, however, was not clear from their results.

The present research, therefore, was designed to investigate how the parts of a temporal pattern interact to determine perceived organization in the total pattern. Our approach to this problem was to determine the organizing principles appropriate to each of the component one-element patterns and then to relate these principles to those found with the two-element patterns from which the one-element patterns were derived.

METHOD

Apparatus. The stimuli were generated from two buzzers which differed in location, quality, and pitch; and from two lights which differed in location and color. The lights were mounted side by side on a vertical board, and appeared 11 in. above a table. They were 1.0 in. in diameter, 2.0 in. apart at the centers. The left light appeared as green, the right as red; each was colored glass transilluminated by an incandescent bulb at a brightness level of .3 ftL. The buzzers were mounted behind the board, 18 in. apart. They both had an intensity of 75-db. sound-pressure level. The left buzzer had a fundamental frequency of 145 Hz., the right 195 Hz.; however, the harmonic differences between the buzzers were more apparent than the frequency differences.

The patterns were generated from a set of interchangeable cams, one cam for each pattern. Each pattern cam had a timing cam such that the duration of each stimulus event was equal to the duration of the blank period between events (a 50%-duty cycle). Electrical switches enabled *E* to select the cam output appropriate to the left and right light and buzzer, and further enabled

---

effect of practice on the perception of temporal patterns, *Quart. J. exp. Psychol.*, in press.

E to turn off either element for the one-element patterns. The cams were driven by a synchronous motor. Mechanical gears were used to hold presentation rate constant (events/sec) across patterns of different lengths.

The *S* and the stimulus display were in a soundproof room, to eliminate equipment noise. An intercom provided two-way communication with *S*, and the call buzzer of the intercom was used to alert *S* for the next pattern. In front of *S*, to the left and right, were mounted two telegraph keys. *S* responded by tapping these in the appropriate sequence. The output of the keys was registered on a Rustrak Continuous Event Recorder. An electromechanical counter registered the number of stimulus events observed by *S*.

Patterns. The *two-element patterns* used in this study were reversals, patterns which when presented in reverse order create new patterns. They are distinguished from patterns which when presented in reverse order remain the same. For instance, the pattern 3112 (with alternative organizations 1123, 1231, and 2311) when presented in reverse becomes 2113 (with alternative organizations 1132, 1321, and 3211). Since 3112 and 2113 are different patterns, they are reversals and form a reversal pair. The pattern 3121 (with alternative organizations 1213, 2131, and 1312), however, is not a reversal, since when it is presented in reverse it becomes 1213 or simply an alternative organization for 3121. The reversal patterns were chosen because they provided a temporal counterbalancing of run lengths for both the two- and one-element patterns.

All possible pairs of reversal patterns of seven, eight, nine, and ten events were used. There was 1 reversal pair that was seven events long; 2 were eight events long; 7 nine events long; and 12 ten events long. Thus a total of 44 two-element patterns was used. These were 3112, 4112, 3122, 4122, 5112, 4113, 3213, 6112, 5122, 4132, 5113, 4123, 3124, 311121, 311112, 212112, 411121, 411112, 312112, 311221, 311122, 311212, and a reversal for each. Of these 44, 26 patterns had four runs and 18 had six runs.

Each of these two-element patterns contained two *one-element patterns*. Since each of the one-element patterns was unique, a total of 88 one-element patterns was used. If a two-element pattern had four runs, then each of its one-element patterns had two runs and two gaps. Likewise, if a two-element pattern had six runs, then each of its one-element patterns had three runs and three gaps. There were 52 patterns with two runs and 36 patterns with three runs.

Subjects. Twenty college students (16 men and 4 women) from the New Haven, Connecticut, area participated. Each was paid nine dollars.

Task. Each *S* served for four days, approximately 70 min. per day. The four experimental sessions were normally completed in one week and never took more than two weeks. Across the four days each *S* was tested on each pattern. All patterns were continuously repeated.

For the two-element patterns *S*'s task was to observe the pattern, wait until *E* stopped the pattern, and then tap the pattern on the two telegraph keys. The task for the one-element patterns was essentially the same except that only one telegraph key was used. For both types of pattern, *S*s were encouraged to sit back and listen to the pattern as opposed to perceiving it quickly, since organization, not difficulty, was of interest. *S*s were allowed to practice tapping while the pattern was still playing. *S* was required to tap the pattern once as soon as possible after its termination, but if he made a mistake he was allowed to tap it again. Average intertrial interval was 25 sec. Each trial was preceded by a warning buzzer. A 10-min. break was provided midway through the session.

Design. On Days 1 and 2 *S* was tested on the two-element patterns. Each pattern was played at the rate of 3 events/sec, since previous work suggested that this rate was probably the easiest for *S*.[4] Each element appeared simultaneously on the appropriate light and buzzer. Thus one element was left light and left buzzer, while the other element was right light and right buzzer. Simultaneous visual and auditory stimuli were used because earlier research had shown that this would be easier for *S* than either modality alone.[5] The aim was to provide *S* with the simplest set of stimulus conditions, since, again, difficulty was not of interest.

The number of repetitions of each two-element pattern was set, on the basis of pilot data, to ensure essentially 100% correct responding. The actual number of pattern repetitions per trial was 20 for patterns of seven or eight events, 25 for patterns of nine events, and 30 for patterns of ten events. These rather lengthy pattern presentations also tend to minimize the effect on perceived organization of the initial starting points of the patterns.[6] In addition, however, initial starting point was counterbalanced across *Ss*, with each pattern starting at each event an equal number of times, or as nearly so as possible.

On both Day 1 and Day 2 *S* saw, in order, 1 two-element pattern seven events long, 2 two-element patterns eight events long, 7 two-element patterns nine events long, and 12 two-element patterns ten events long, for a total of 22 patterns. Order of presentation within the seven-, eight-, nine-, and ten-event patterns was counterbalanced with the restriction that both patterns of a reversal pair were not seen on the same day.

Left light and left buzzer (L) versus right light and right buzzer (R) were counterbalanced within and between patterns. Within patterns, half of the *Ss* experienced the pattern in one form, and half in its complementary form. For example, the pattern 3112 appeared as RRRLRLL . . . for half of the *Ss* and as LLLRLRR . . . for the other half. Between patterns, each *S* saw an equal number of red and green lights across Days 1 and 2.

Thus for each *S* all data on two-element patterns were obtained before *S* experienced one-element patterns. Our reason for this experimental arrangement was that we did not want to predispose *S* to attempt to deal with the two-element patterns as one-element patterns. It was not until Days 3 and 4, then, that *S* was tested on the one-element patterns. Each of these patterns was played at the rate of 3 events/sec. For half of the *Ss* (assigned randomly), all of the one-element patterns were played on the right light and right buzzer; while the other half of the *Ss* observed only the left light and left buzzer.

Since each two-element pattern produced two one-element patterns of the same number of events, on Day 3 and on Day 4 *S* observed, in order, 2 one-element patterns based on seven events, 4 based on eight, 14 based on nine and 24 based on ten. Within a given length (seven, eight, nine, or ten events), order of presentation was counterbalanced with the restriction that both one-element patterns from a single two-element pattern not be seen on the same day. Since each two-element pattern had a reversal, of necessity each one-element pattern also had a reversal. Therefore, order of presentation was further restricted in that a given one-element pattern and its reversal were not presented on the same day.

---

[4] Garner and Gottwald (see n. 3 above).
[5] Handel and Lewis (see n. 3 above).
[6] Garner and Gottwald (see n. 3).

All one-element patterns were presented for 10 complete repetitions. Pilot data suggested that this was an adequate trial length regardless of a given pattern's two-element base. Each pattern was started an equal number of times at each positive event, or as nearly so as possible.

Basic data. The actual individual recorded datum was the pattern tapped out by *S* on the one or two telegraph keys. This response was considered a description of the pattern as perceived by *S*, exactly analogous to verbal descriptions used in other experiments on this topic. Each pattern, however, had a fixed number of alternative organizations, so the description was considered *S*'s use or choice of one of the available organizations. The summary measure for each pattern was then the frequency of use of each of the available alternative organizations.

If statistical tests were carried out, the appropriate sampling variable would thus be the pattern, not the individual *S*, since basically the pertinent experimental questions concerned the generality of a result across all patterns having certain common characteristics. We present no significance tests, however, because the validity of the argument rests in the consistency of a set of explanatory principles across the several sets of data presented—on one- and two-element patterns, and on patterns of different numbers of runs. Still further, these same principles can be shown to explain previous data on the distribution of pattern descriptions, as well as data on difficulty of pattern perception and learning.

RESULTS: ONE-ELEMENT PATTERNS

The results from the one-element patterns are presented first, and from these there emerge two principles of perceptual organization. The results from the two-element patterns are then considered as a function of the organizing principles of the component one-element patterns.

Preliminary analyses established two things. First, errors were too few to require analysis. Of a total of 1,760 one-element pattern responses, only 15 were incorrect.[7] Second, pattern descriptions did not break runs of events. Thus, as with two-element patterns, the run was the psychologically meaningful unit, and further analyses here need only consider runs.

---

[7] An error was recorded when the response did not contain the appropriate numbers of positive events in each run and the appropriate gaps. However, no attempt was made to distinguish the relative lengths of the gaps. Furthermore, since it is impossible to determine whether a pattern description started with a gap, all descriptions were scored as though they started with a run of positive events. Thus for patterns with two runs and two gaps, a correct description needed only to show two runs of the correct number

Patterns with Two Runs

The one-element, two-run patterns presented the subject with just two possible perceptual organizations, since runs were not broken and there were only two runs. Which of these two runs would be used to start the pattern could depend on two obvious properties of the patterns: differences in length of the two gaps, and differences in length of the two runs. Both factors operate to provide principles of perceptual organization.

Gap principle. Across all one-element patterns having two runs, the subjects used organizations in which the longer gap was at the end 79% of the time and was in the middle the remaining 21% of the time. According to the principle implicit in these results, the pattern organization 1′3″ would be chosen over 3″1′, since the continuous sequence is most often experienced as repeated pairs of runs separated by the longer gap.

Run principle. Also, for all of these patterns, the subjects chose organizations in which the longer of the two runs was the first run 61% of the time and was the second run the remaining 39% of the time. By this principle, the pattern organization 3″1′ would be chosen over 1′3″.

Compatibility of organizing principles. The fact that each of these organizing principles was used over half the time makes it clear that the two principles did not exist in these patterns as mutually exclusive alternatives. The fact that the two percentages of choice were not identical also makes it clear that the two principles were not perfectly correlated over this set of patterns. In order to further clarify the role of these two principles, we need to distinguish between patterns in which the gap and run principles were compatible (would lead to the same organization) and patterns in which they were incompatible (would lead to different organizations). The pattern 3′1″ is an example of a pattern with compatible

---

of positive events. For patterns with three runs and three gaps, a correct description needed to have the three runs of the correct length and in a correct order.

principles in that either principle leads to this organization. The reversal of this pattern is one in which the organizing principles are incompatible, since the run principle predicts the organization 3″1′, while the gap principle predicts the organization 1′3″.

When two patterns are reversals of each other, one of the pair of patterns contains compatible principles, while the other contains incompatible principles. The use of reversal pairs ensured this contrast in the present experiment, and since we used 52 patterns with one element and two runs, 26 of them contained compatible principles and 26 contained incompatible principles. Table I presents the data separately for these compatible and incompatible patterns. For the 26 compatible patterns, the organization predicted by both the gap and the run principles was chosen 90% of the time. Thus when both of these principles would predict the same organization, that organization was nearly always chosen.

When the two organizing principles would predict different organizations, then the organization predicted by the gap principle was chosen 68% of the time. Thus it appears that for these patterns the gap principle tended to dominate the run principle. This dominance can be seen also by examining the incompatible patterns in more detail, as shown in Table I. When two organizing principles are available, the choice should be influenced by the relative strength of the principles. Strength is indicated in Table I in terms of the relative difference in the lengths of the runs and the lengths of

TABLE I

PERCENT USE OF ALTERNATIVE ORGANIZATIONS FOR ONE-ELEMENT, TWO-RUN PATTERNS FOR COMPATIBLE AND INCOMPATIBLE ORGANIZING PRINCIPLES

	Organization		N
Pattern type	Gap & Run	Other	(Patterns)
Compatible principles	90	10	26
	Gap	Run	
Incompatible principles			
Gap difference > run difference	86	14	10
Gap difference = run difference	71	29	6
Gap difference < run difference	49	51	10
Overall	68	32	26

Note: For compatible patterns, the organizing principles of gap and run predict the same organization. For incompatible, they predict different organizations.

the gaps. When the gap difference was greater than the run difference (e.g., 1′2‴), then the organization predicted by the gap principle was used 86% of the time, nearly as often as when the principles were compatible. When the gap difference was equal to the run difference (e.g., 2′3″), the organization predicted by the gap principle was used 71% of the time. Only when the gap difference was less than the run difference (e.g., 4″2′) was the run principle used more than the gap principle, and even then just barely, with 51% of the choices being those predicted by the run principle.

Thus two simple principles determine pattern organization: have the largest gap at the end, separating successive patterns in the ongoing sequence; and have the longest run at the beginning. Of these two, the gap principle was clearly stronger in this set of patterns. Variability of pattern organization was determined by the interaction of these two principles, being very low when the principles were compatible or when the gap principle dominated, and becoming greater when the strength of the run principle was equal to or greater than that of the gap principle.

Patterns with Three Runs

The descriptions of one-element, three-run patterns continued to show the primary effects of the run and gap principles, but with these more complex patterns the picture is more complex. There may also have been an additional organizing principle operating.

Overall analyses show that in 28 of the 36 patterns there was one gap longer than the remaining two gaps, and 81% of the pattern descriptions had this longest gap at the end, separating the successive patterns in the continuing sequence. There were also 28 patterns (20 being the same as above) in which one run was longer than the other two, and for these patterns 43% of the pattern descriptions (chance being 33%) were with the longest run at the beginning.

With each of these organizing principles, eight of the patterns did not have an unambiguous prediction because at least two runs or gaps were equally 'longest' (1′2′2″, where two runs have a length of two; or 3′2′1′, where all three gaps are of the same length). For the eight patterns in which use of the run principle would make an ambiguous prediction but the gap principle would make an unambiguous prediction, the organization used was predicted by the gap principle 88% of the time. For the eight patterns in which use of the gap principle would lead to an ambiguous prediction but the run principle would make an unambiguous prediction, the organization was predicted by the run principle 44% of the time. However, there was a suggestion of an additional principle operating here, which is that the sequence of run lengths produces an upward progression. In two of these eight patterns (1′2′3′, 1′2′4′) a simple upward progression existed, which, of course, gives an organization different from one starting with the longest run. For these two patterns the upward progression was used a total of 21 times compared to 18 for the organizations starting with the longest run. There were too few patterns contrasting these principles to allow great confidence in the importance of progression. Furthermore, we can only separate the progression principle from the run principle with upward progressions, since with downward progressions of run lengths the run principle alone would predict the organization chosen.

Table II shows a more detailed analysis of data for the 20 patterns in which both the run principle and the gap principle lead to unambiguous predictions. For the eight patterns in which these two principles were compatible and would thus lead to a single predicted organization, that organization was used 81% of the time. For the other 12 patterns, in which the two principles were incompatible and would thus lead to different predictions, the gap principle predominated, although as its strength decreased there was an

TABLE II

PERCENT USE OF ALTERNATIVE ORGANIZATIONS FOR ONE-ELEMENT, THREE-RUN PATTERNS FOR COMPATIBLE AND INCOMPATIBLE ORGANIZING PRINCIPLES

Pattern type	Organization			N (Patterns)
	Gap & Run		Other	
Compatible principles	81		19	8
	Gap	Run	Other	
Incompatible principles				
Gap difference > run difference†	95	5	0	2
Gap difference = run difference†	87	8	5	3
Gap difference = run difference	65	24	11	5
Gap difference < run difference†	72	28	0	2

† Organizing by gap also provides upward progression.

increased use of the run principle. However, for three of the groups of patterns, the use of the gap principle also allowed use of an upward progression (e.g., 1′1′2‴), and the use of the organization predicted by the gap principle was greater. This fact suggests that the principle of upward progression was adding its effect to the gap principle. Thus these data show that the gap and run principles found to determine perceptual organization with one-element patterns of two runs also determined perceptual organization and variability of organization when there were three runs.

Summary

The results with two- and with three-run patterns lead to the following generalizations about the operation of the run and gap principles:

1. If the run and gap principles are compatible and thus predict the use of a single organization, that organization is nearly always used.
2. If the run and gap principles are incompatible and thus predict the use of different organizations, two different organizations are used to some extent. Thus incompatibility of the run and gap principles leads to greater variability of use of organizations.
3. When the run and gap principles operate incompatibly, the gap principle seems to be stronger. This result, however, is probably most susceptible to change with variations in experimental procedure. For example, with the 50%-duty cycle used in these experiments, a gap nominally of the same length as a run was actually longer, because the positive element did not occupy all of the time available for it.

RESULTS: TWO-ELEMENT PATTERNS

There have been no previous studies on perception of temporal patterns using patterns of seven or ten events. The patterns of these lengths used in the present study showed no principles of organization different from those established for patterns of eight and nine events. This result is not surprising since runs, not individual events, are the meaningful units, and both four- and six-run patterns have previously been used. Of the 880 responses to the two-element patterns, 49 were in error and 5 were correct but broke a run of identical elements. These responses were too few to allow determination of any systematic trends and were dropped from further analysis.

Overall Analysis

The primary purpose of this paper is to determine how the parts of two-element temporal patterns combine to influence perceived organization of the total pattern. The simplest prediction rule for two-element data would be that they are simply a composite or summation of the component one-element data. In order for this summation to be meaningful, however, we must first establish the structural parallels between the two types of patterns.

Patterns with four runs. There were 26 patterns having four runs. With four runs there are four possible organizations, and each organization can be specified in terms of its appropriate one-element principle or principles. For each of these patterns one of the component one-element patterns had compatible run and gap principles, while the other had incompatible run and gap principles. (This relation is a consequence of the use of reversal patterns.) Thus, of the four runs that could be used to start the pattern, one would be predicted by gap and run (compatible one-element pattern), the next run by gap (incompatible pattern), the next by nothing (alternative run of compatible pattern), and the last by run (incompatible pattern). For instance, the alternative organizations for 4112 would be predicted, in order, by gap and run (4112), gap (1124), nothing (1241), and run (2411).

Table III presents overall data, organized in this fashion, for all four-run, two-element patterns. In addition it presents data formed from a composite of the component one-element patterns,

TABLE III

TOTAL PERCENT USE OF ALTERNATIVE ORGANIZATIONS OF FOUR-RUN, TWO-ELEMENT PATTERNS COMPARED WITH THE DISTRIBUTION OBTAINED FROM A COMPOSITE OF THE COMPONENT ONE-ELEMENT PATTERNS

	Organization			
Pattern type	Gap & Run	Gap	Other	Run
Composite one-element†	45	34	5	16
Two-element	39	27	7	27

† These values are direct averages of the data shown in Table I.

obtained by simply averaging the overall figures for the appropriate pattern alternatives in Table I. From this table it can be seen that the first conclusion for the one-element patterns was verified in the data for the two-element patterns: the organization most frequently used (39%) was the one starting with the run, in the compatible one-element pattern, predicted by both the run and gap principles. The second conclusion for the one-element data was that when these two principles are incompatible, both organizations are used to some extent. This conclusion was also confirmed in the two-element data (27% and 27%). The third conclusion was that when the gap and the run principles are incompatible, the gap principle tends to be the stronger. This conclusion was not verified, since the organization predicted by gap and that predicted by run were used equally often in the two-element data. Since this was the weakest of the three conclusions, its disconfirmation was not considered critical.

Patterns with six runs. An overall analysis and comparison like the one for four-run patterns was not possible with six-run, two-element patterns. The greater complexity of the latter precludes any single organizational arrangement into which all of the patterns can be put. These patterns are analyzed by pattern type and preferred element, below.

Analysis by Pattern Type and Preferred Element

Our demonstration of overall agreement between the use of different organizations in two-element data and uses of the equivalent organizations in one-element data is of limited significance, however, since the analyses in both Tables I and II clearly show differences in the distribution of use with different specific subtypes of pattern. We need to know whether the two-element data conformed *in detail* to the one-element data.

We also need to know whether an individual subject observing a particular two-element pattern considers alternative organizations from both elements or whether he primarily considers organizations appropriate to just one of the elements. If the latter, greater prediction would be obtained by knowing to which of the elements the subject was attending, or, as is argued below, which of the elements he had taken as 'figure.' An analysis of the use of each of the two stimulus elements showed that across all subjects and patterns, 51% of the descriptions began with the left element and 49% began with the right element. Since the patterns presented were counterbalanced in terms of right versus left element, there was no overall preference for one element over the other. However, the variance between subjects with respect to preference for starting pattern descriptions with the two elements was, indeed, highly significant ($\chi^2 = 216.95$, $df = 19$, $p < .001$). Some subjects consistently began descriptions with the left element, while others began with the right element.

Therefore, the subjects were divided into two groups: those who preferred the left element and those who preferred the right element. There were eight subjects in each of these groups, each subject using his preferred element to start the description at least 26 times out of the total of the 44 descriptions given. The remaining four subjects were deleted from further analysis. With this information on subjects' preferred elements, further analyses were carried out to investigate the effect of particular type of pattern and the effect of individual preferences for stimulus elements.

Patterns with four runs. Since each four-run, two-element pattern contained one compatible and one incompatible one-element pattern, and since left versus right element was counterbalanced, the preferred element could provide either a compatible or incompatible one-element pattern. The two-element pattern descriptions used for these two possibilities are shown in Table IV. This table is like Table I for one-element patterns, since in effect we are looking at the preferred element in terms of its one-element pattern characteristics. In fact, however, the patterns consisted of two elements; therefore the organization can shift to a description appropriate to the nonpreferred element.

Table IV shows that when the preferred element was a compatible one-element pattern, subjects used descriptions starting with the preferred element at the point predicted by both gap and run principles 61% of the time, and at the point predicted by neither principle (but still with the preferred element) only 10% of the time. The remaining organizations used as descriptions represent shifts to the other element (a necessarily incompatible pattern), with 13% of the descriptions at the point predicted by gap and 16% at the point predicted by run.

When the preferred element produced a one-element pattern with incompatible gap and run predictions, the results varied as a

TABLE IV

PERCENT USE OF ALTERNATIVE ORGANIZATIONS FOR TWO-ELEMENT, FOUR-RUN PATTERNS AS A FUNCTION OF ONE-ELEMENT ORGANIZING PRINCIPLES IN THE PREFERRED ELEMENT

Pattern type	Organization				N
	Gap & Run	Gap	Other	Run	(Patterns)
Compatible principles	61†	13	10†	16	26
Incompatible principles					
Gap difference > run difference	22	58†	6	14†	10
Gap difference = run difference	23	40†	10	27†	6
Gap difference < run difference	17	24†	1	58†	10
Overall	20	41	5	34	26

† Subject begins organization with preferred element.

function of the relation between the principles, as shown in the lower part of Table IV. These results show effects very similar to those shown for one-element patterns in Table I, with the gap principle being the better predictor when differences in gap length were greater than or equal to differences in run length, and with the run principle being the better predictor only when it was the stronger principle. Shifts to the nonpreferred element (a necessarily compatible pattern) accounted for the remaining organizations, with 20% at the point predicted by gap and run, and 5% at the point predicted by neither principle.

The results with the four-run patterns show that use of available organizations of two-element patterns was quite predictable from the results of the component one-element patterns. Further, given the element which is preferred or salient, the two-element organization was that which starts with the preferred element, and starts with that run in the preferred element which would have been used if only the one-element pattern were experienced. In addition, variability of organization was greater when incompatible principles were present in the preferred element.

Patterns with six runs. Because one-element patterns with three runs may have compatible and incompatible run and gap predictions as well as ambiguous run or gap predictions, prediction for

TABLE V

PERCENT USE OF ALTERNATIVE ORGANIZATIONS FOR TWO-ELEMENT, SIX-RUN PATTERNS AS A FUNCTION OF ONE-ELEMENT ORGANIZING PRINCIPLES IN THE PREFERRED ELEMENT

Pattern type	Organization				N
	Gap & Run	Gap	Other	Run	(Patterns)
Gap and run compatible	70†	7	8†	15	8
Gap and run incompatible	19	49†	2	30†	8
Gap with run ambiguous		61†	9	30	8
Run with gap ambiguous		14	22	64†	8

† Subject begins organization with preferred element.

the two-element, six-run patterns involved more types of pattern, most of which are summarized in Table V. Eight patterns had one component one-element pattern in which the gap and run principles were compatible. When that element was preferred, the organization chosen was the compatible gap and run organization 70% of the time, as shown in the top line of data. The other one-element patterns for these eight patterns had gap and run principles incompatible, and when this element was preferred gap was used as the organizing principle 49% of the time and run 30% of the time. There were also two patterns for which both component one-element patterns had incompatible predictions. These data (based on only 30 descriptions) do not appear in Table V, but showed the same pattern of results: 47% of the organizations were predicted by the run principle and 43% by the gap principle.

Eight of the patterns had component one-element patterns in which the gap principle would give an unambiguous prediction with the run principle giving an ambiguous prediction (e.g., $1'2'2''$) or vice versa (e.g., $3'1'2'$). These cases are shown on the bottom two lines of Table V. Whichever principle would lead to an unambiguous prediction was used over 60% of the time.

There was no suggestion in these data that an upward progression plays any role in determining perceptual organization, as it appeared to do with one-element patterns. These results with six-run patterns do, however, show that perceptual organization of two-element patterns can be understood from knowledge of the organizing principles pertinent to the component one-element patterns, as was found for four-run patterns.

DISCUSSION

The results for one-element patterns show that perceptual organization was determined by two primary principles: *gap,* in which the pattern is organized so that the largest gap occurs at the end of the pattern; and *run,* in which the pattern is organized so that it starts with the longest run. These results also show that variability of perceptual organization was determined by compatibility: when the two principles are *compatible* in a particular pattern, a single organization is perceived or used; when the two principles are *incompatible,* more than one organization is used. Thus which organizations were used was determined by the run and gap principles, and how many organizations were used depended on the relations between these two principles in a given pattern.

Figure-Ground Relations

Since the pattern organizations used with two-element patterns were composites of the pattern organizations used with the two one-element patterns composing those two-element patterns, it seems reasonable to suppose that the two-element pattern was perceived with one element as figure and the other element as ground. The fact that the subjects in the present experiment showed strong preferences for using organizations appropriate to one element rather than the other argues for this figure-ground approach to the problem.

If these patterns are perceived as figure-ground, the perceived organization of a two-element pattern is determined by the organization of the one-element pattern that is figure. But does structure of the one-element patterns determine figure-ground relations, or do figure-ground relations determine perceived structure?

Structure of one-element patterns. The data on pattern organization for two-element patterns show that both the number of organizations used as well as the particular organizations used depended on the organizational properties of the two one-element patterns composing the two-element pattern. However, the relationship was not one of structural interaction in the sense that the two one-element patterns interacted to produce a new set of organizations. To illustrate, consider two-element patterns of four runs. In each such pattern used in the present experiment, one of the component one-element patterns had the run and gap principles compatible, while the other one-element pattern had them incompatible. It seems quite possible that the two-element pattern might have been organized in such a way that the compatible one-element pattern became figure, while the incompatible one-element pattern became ground. If this had occurred, then the two-element patterns with high variability of organization would be those in which each of the one-element patterns had a strong (thus competitive) organization, while the two-element patterns with low variability of organization would be those with one of the one-element patterns having much stronger organization than the other. The results do not show any such effect. Furthermore, the two-element, four-run data (Table IV) show no greater tendency for subjects to shift from their preferred element when the preferred element provided a pattern with incompatible principles than when the preferred element provided a pattern with compatible principles (26% versus 29%). Likewise, for those six-run patterns in which an equivalent shift was possible (top two data lines in Table V), there was no greater tendency to shift toward patterns with compatible principles than to shift toward patterns with incompatible principles (22% versus 21%).

Properties of elements. Thus structural properties such as compatibility of organizing principles did not determine which element of a pattern was figure and which was ground. Rather, which element was figure was determined first, and then the organizations perceived were those available in the one-element pattern that was the figure. Since the available perceptual organizations were not used equally often in the two one-element patterns, then the factors determining which element was figure and which ground also determined the perceived organizations, as well as the variability of organizations. Since previous research (see n. 2) has shown that difficulty of pattern perception is directly related to variability of organization, then these same factors also affected difficulty of pattern perception.

In the present experiment, no experimental manipulations were used to affect figure-ground relations of the elements; the fact that equal numbers of subjects used each of the two elements as figure was simply happenstance. In a previous experiment (see n. 1), however, it was found that the higher-pitched of two tones was more frequently used as figure than as ground. Therefore such an experimental manipulation—as well as others pertinent to visual stimulus patterns—could be used to alter figure-ground relations, and thus to alter the perceived organization and difficulty of two-element temporal patterns.

Pattern Length and Number of Runs

Research on both the perception of more rapidly presented temporal patterns and the learning of slower patterns[8] has shown the run rather than the single event to be the pertinent psychological unit. This fact means that length per se of a temporal pattern is not a crucial variable. But neither does it mean simply that number of runs is the crucial variable. Rather, these results show that interrelations between run lengths determine pattern complexity and difficulty, in ways quite predictable from a small number of organizing principles.

To illustrate the lack of importance of both total pattern length and number of runs per se, as well as the importance of the configuration of pattern lengths, consider the following rather extreme example. Although there are no data on the pattern of seven events 211111, it is clear from the present analysis that it should be an easy pattern to perceive, since each component one-element pattern has either the run or the gap principle operating unambiguously, with the other principle being ambiguous. Thus there is just one acceptable organization for either of the one-element patterns, so that regardless of which element is figure, the pattern should be an easy one. If we increase the length to eight events by increasing the one longest run, to give 311111, the pattern should be at least as easy, and probably easier, because either the gap or run principle can operate more strongly. Previous data from Royer and Garner (see n. 2) show that 21.4 events were heard, on the average, before accurate tracking of the pattern could occur.

If the pattern length is increased by making the pattern 211112, the pattern should be more difficult, because while one of the component patterns has the run and gap principles compatible (2′1′1″), the other component pattern has them incompatible (1′1′2″) with the gap and run principles being equally strong in terms of differences of lengths. This incompatibility of one of the component patterns should make the organization more difficult to perceive; and, in fact, the previous research showed 30.5 events were heard before tracking began.

---

[8] See, for example, F. Restle, Grammatical analysis of the prediction of binary events, *J. verb. Learn. verb. Behav.*, 6, 1967, 17–25; P. C. Vitz, Information, run structure, and binary pattern complexity, *Percept. & Psychophys.*, 3, 1968, 275–280.

Still further, if the pattern is increased in length by making it 212111, it should be even more difficult, since now neither component pattern provides an unambiguous organization with either organizing principle (2′2′1′ and 1″1′1″). The previous research showed that 48.8 events were heard before tracking began. And as a last example, if the pattern of eight events is 211211, difficulty should become very great, because both component one-element patterns have incompatible run and gap principles operating (2′1″1′ and 1′2′1″). The previous research showed that 103.9 events were heard before tracking could begin. These four examples are the only patterns which can be produced by adding one event to the original pattern. Difficulty of pattern perception for these four patterns ranged from among the easiest to the most difficult.

Thus, clearly the critical factor in determining number of alternative organizations and difficulty of pattern perception is neither length of the pattern nor number of runs. Rather it is the interaction of principles of temporal pattern perception in the two component one-element patterns. It is for such reasons that we argue that perception of these patterns is based on wholistic principles—that all elements and runs interact to produce the perceived organization, and to make the pattern an easy or difficult one in which to perceive organization.

MAINTENANCE OF FIRST-LIST ASSOCIATIONS DURING TRANSFER

LEO POSTMAN and JANAT FRASER PARKER,
University of California, Berkeley

Abstract. This study investigated the effects on transfer and retroactive inhibition of first-list recall during the test trials of second-list learning (simultaneous method). In Experiment I this procedure was used under four paradigms of transfer: A–B, C–D; A–B, A–C; A–B, A–Br; and A–B, A–B′. In Experiment II a direct comparison was made for two paradigms (C–D and A–C) between the simultaneous method and the conventional procedure (successive method). In the simultaneous condition the differences among paradigms in transfer learning paralleled those obtained under the successive treatment. The maintenance of first-list responses resulted in only a moderate retardation of transfer performance. Retroactive inhibition was reduced but not eliminated when the first list continued to be recalled in the transfer stage. In the simultaneous situation the amounts of retroactive inhibition for C–D and A–C did not differ reliably and appeared to reflect primarily a reduction in the availability of first-list responses. Progressive unlearning of first-list associations was in evidence only for A–Br. Mediated facilitation was observed under the A–B′ paradigm.

The conventional experimental arrangements used in the study of negative transfer and interference may be regarded as paradigms of habit substitution. Consider, for example, the classical condition of negative transfer in which different unrelated responses are attached successively to the same stimuli (A–B, A–C). The task presented in the transfer stage leads the subject to substitute C for B, rather than to add C to B as an alternative response to A. Response competition (the erroneous elicitation of B during the acquisition of A–C) serves to reinforce the subject's disposition to adopt this mode of attack on the transfer task. Substitution implies at a minimum that the old responses fail to be rehearsed; it is likely that they are actively suppressed.[1] When recall of the first list is required, the process of substitution must be reversed abruptly.

To what extent is the paradigm of substitution representative of conditions of interference outside the laboratory? Many verbal tasks are completed and not practiced again, and hence are subject

---

Received for publication November 3, 1969. This research was supported by Grant MH-12006 from the National Institute of Mental Health.

[1] For a discussion of the possible role of response suppression in retroactive inhibition, see L. Postman, K. Stark, and J. Fraser, Temporal changes in interference, *J. verb. Learn. verb. Behav.*, 7, 1968, 672–694.

to cumulative interferences from other activities. It is difficult, however, to think of many instances in which the old habits become explicitly wrong when the new ones must be acquired. More likely than not, the old associations remain correct as far as they go and may, indeed, continue to be used in the appropriate context. Learning the names of objects in different languages provides a common example. Thus, the A–B, A–C paradigm applies to many situations in which there is an accretion of associations to the same stimuli rather than the substitution of the most recent responses for the earlier ones. The same considerations apply to such other paradigms of negative transfer as A–B, A–Br.

It is reasonable to suppose that the amount of retroactive inhibition is less under conditions of habit accretion than of habit substitution.[2] Continued use serves to reinforce the older habits and increase their resistance to interference. Moreover, if the alternative associations to a stimulus are activated intermittently, the suppression of early responses in favor of more recent ones, if it does occur, can be only transitory. The question then becomes one of the conditions under which, and the extent to which, retroactive inhibition does, in fact, develop when there is an accretion of successive habits. Does the establishment of new associations necessarily entail some reduction in the availability of the previously learned ones, even though the latter remain correct and continue to be used? If so, what is the temporal course of the retroactive losses as the new habits grow in strength? One possibility is that some extinction or suppression of competing prior responses is inevitable during the initial stages of the acquisition of the new associations. That is, there may be an initial phase of habit substitution even when the successive associations are cumulative rather than mutually exclusive. The process of unlearning is presumably not sustained, however, once both the old and the new habits begin to be activated intermittently. On this assumption, retroactive inhibition should set in rapidly but level off early in the transfer stage.

A further question of interest is whether differences in the amount of retroactive inhibition among paradigms of specific transfer show a parallel pattern under conditions of accretion and of substitution. In the latter case the amount of retroactive inhibition is, in general, related directly to the degree of associative interference. This relation should become less pronounced when old responses occurring in the transfer stage are no longer treated as errors that must be eliminated altogether; under these circumstances the frequency of explicit and implicit interlist intrusions may have only a limited influence on the level of unlearning. Under all paradigms the subject's task is to differentiate old from new responses while maintaining both for subsequent use. Thus, the differences among paradigms may be expected to diminish as the effective definition of the transfer task changes from one of substitution to one of accretion.

---

[2] This expectation receives some support from a finding reported by R. Popp and J. F. Voss, Aquisition: Transfer and retention of S–R_1 and S–R_2 associations as a function of S–R_1 and S–R_2 pattern, *J. exp. Psychol.*, 70, 1965, 304–308. When two different responses are learned to the same stimulus according to a pattern of alternation, unlearning decreases as the shifts from one response to the other become more frequent.

In terms of experimental operations, there are two departures from the conventional procedures of transfer studies that serve to define a condition of accretion rather than of substitution. First, the subject is informed prior to the beginning of the transfer task that the first-list associations will have to be recalled at a later time.[3] Second, tests of first-list recall are, in fact, interspersed among the trials of transfer learning. The operations defining a situation of accretion must specify the schedule of such tests. The number and the temporal distribution of the occasions on which the original responses are reactivated constitute a critical difference between the two kinds of transfer learning. In the present study, tests of recall of the old and of the new associations occurred with equal frequency and in close temporal proximity. Specifically, recall of the first-list as well as of the second-list responses was required on every test trial in the transfer stage. This schedule was considered optimal in an initial investigation of interference under conditions of accretion, since it would prevent persistent dominance of either the first-list or the second-list responses. In Experiment I this procedure was used to measure the course of changes in first-list recall under different paradigms of specific transfer. A direct comparison of both transfer and retroactive effects under conditions of substitution and of accretion was made for selected paradigms in Experiment II.

---

[3] Informing subjects prior to interpolated learning that there will be a test of first-list recall has been shown to reduce the amount of retroactive inhibition. See L. Postman and K. Stark, Retroactive inhibition as a function of set during the interpolated task, *J. verb. Learn. verb. Behav.*, 1, 1962, 304–311.

EXPERIMENT I

Method

Design. The experimental design comprised four transfer groups and two control groups. The four paradigms of transfer were: A–B, C–D; A–B, A–C; A–B, A–Br; A–B, A–B′. Thus, in addition to the reference condition (C–D), there were two paradigms of negative transfer (A–C and A–Br) and one of positive transfer (A–B′). In the positive case retroactive facilitation was expected, since the acquisition of the second-list response should provide an opportunity for mediation, or parasitic reinforcement. The control groups learned and recalled a single list. The reasons for the inclusion of two control groups, which were given recall tests of different durations, are detailed below, in the section on procedure.

Lists. The learning materials were lists of 12 paired associates, with single letters as stimuli and two-syllable adjectives as responses. There were two sets of stimulus letters and eight sets of response words, which were used to generate 16 different lists. The first and second lists were fully counterbalanced. Under all transfer conditions each of the 16 lists was used once as the first task and once as the second task. The lists were also used once each in the control condition.

Within a list each response word began with a different letter. The stimulus letter and the initial letter of the response were never the same. Furthermore, different responses to the same letter (as in the A–C paradigm) did not begin with the same letter. Meaningful response similarity was minimized, except as required in the A–B′ paradigm. The B and B′ terms were synonyms chosen from Haagen's tables and from norms obtained in a prior standardization procedure.[4]

Procedure. The first list was learned to a criterion of 7/12 by the study-test method. The pairs were presented in four different orders on the study trials, and four other orders were used on the test trials. The rate of presentation on all trials was 2 sec., with a 4-sec. intertrial interval.

The interval between the first and the second list was 3 min. There were six study-test cycles of second-list learning. On the study trials the pairs were presented at a 2-sec. rate. There were four different orders of presentation. The method of modified free recall (MMFR) was used on the test trials. The *S*s in the transfer groups were given 2 min. in which to write down both the first-list and the second-list responses to the appropriate stimuli. In the A–C, A–Br, and A–B′ conditions, there were two blanks next to each stimulus letter on the test sheet. In the case of the C–D paradigm, the stimuli from both lists were included, in separate columns, with a single blank beside each. The stimuli were listed on the test sheets in four different orders.

The control *S*s were given six test trials on the original list. These trials were spaced so as to begin at the same time as the corresponding tests in the transfer conditions. The intervals between successive tests were filled with a digit-symbol substitution task. One control group was allowed 1 min. for the completion of each test, and a second control group was given 2 min. per test. These two groups are designated Control 1 and Control 2 respectively. The use of a 1-min. test was designed to take account of the fact that the control *S*s had to recall half as many responses as the *S*s in the transfer groups. The group given 2-min. trials was added in order to equate precisely the temporal distribution of tests in the control and the transfer conditions. The estimate of the absolute amount of retroactive inhibition was, of course, maximal with respect to Control 2.

---

[4] The standardization procedure is described in L. Postman, Differences between unmixed and mixed transfer designs as a function of paradigm, *J. verb. Learn. verb. Behav.*, 5, 1966, 240–248.

The *S*s were given full instructions at the beginning of the transfer phase about the procedure to be followed on the test trials. They were informed that they would be free to write down the responses in any order they wished and were encouraged to guess. A new test sheet was handed to *S* at the beginning of every test trial and removed at the end of the recall interval.

Subjects. There were 16 *S*s in each of the six groups. The *S*s were undergraduate students at the University of California who were not necessarily naïve to verbal-learning experiments. Assignment to conditions was in blocks of six, with one *S* from each condition per block. The running order within blocks was determined by means of a table of random numbers, as was the assignment of *S*s within groups to particular combinations of lists.

Results

First-list learning. The mean number of trials to a criterion of 7/12 on the first list was 4.78, with the means of individual groups ranging from 3.94 to 6.12. The mean number correct on the criterial trial was 7.91, with a range from 7.81 to 8.06. There was no significant variation among groups for either measure (both Fs < 1).

Second-list learning. The mean numbers of correct second-list responses given on the successive MMFR tests in the transfer phase are shown in Figure 1. The stringent scores based on responses to the appropriate stimuli are plotted. The differences among paradigms conformed to the conventional pattern. Relative to the C-D reference condition, there was negative transfer for both A–C and A–Br; the effect was more pronounced for the latter. There was facilitation for A–B′. The mean total numbers of correct responses given on the six test trials were: for C–D, 50.2; for A–C, 43.6; for A–Br, 40.6; and for A–B′, 55.6. The differences among paradigms were significant [F (3, 60) = 5.79, $p < .01$]. The conclusions remain the same when only the first test

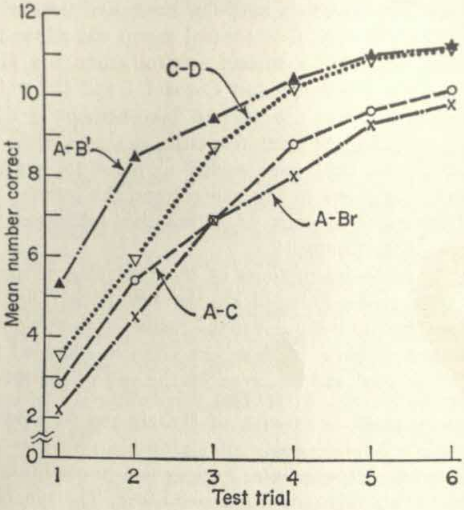


Fig. 1. Mean numbers of second-list responses given correctly on each test trial in Experiment I

trial is considered. In general, then, the trends in transfer were comparable to those obtained under conventional conditions.

The differences between stringent and lenient scores can be summarized briefly. The latter were based on all responses from the list, regardless of whether or not they were correctly associated. With the exception of the A–Br paradigm, the lenient scores exceeded the stringent ones only by small amounts. For A–Br the lenient scores were uniformly high because the responses were the same in the two lists.

First-list recall. The stringent measures of first-list recall for the successive MMFR tests are shown in Figure 2. The mean total numbers correctly recalled on the six tests are presented in Table I. Attention is called first of all to the fact that the performance of Control 2 exceeded that of Control 1. The separation between the two conditions remained constant over trials; on the average, the extension of the test period from 1 to 2 min. permitted the recall of approximately one additional item. Thus, Control 2 yielded larger estimates of the amounts of retroactive inhibition than did Control 1. However, the difference between the mean total scores of the two groups was not significant [t (30) = 1.47].

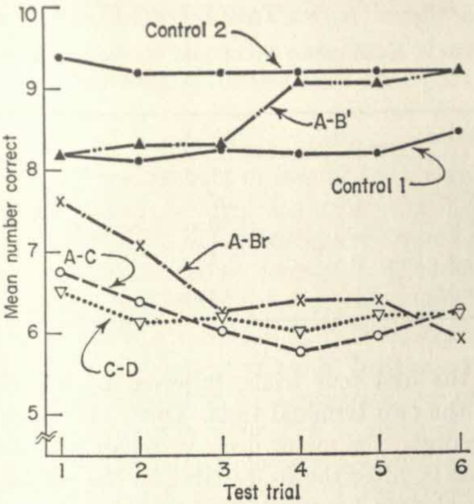


FIG. 2. Mean numbers of first-list responses recalled on each test trial in Experiment I

The level of recall, as indexed by the mean total number recalled, was lower for the C–D, A–C, and A–Br paradigms than for the control conditions. As Table I shows, the overall amounts of retroactive inhibition were closely comparable for these three conditions. The A–B′ paradigm fell between the two control treatments. The scores of Control 1 and of the four transfer groups were used in a first analysis of variance of the total numbers correct. The variation among the transfer groups was significant [$F(3, 75) = 9.12, p < .01$].

Orthogonal comparisons showed that A–B′ yielded significantly higher recall scores than the remaining transfer paradigms [$F(1, 75) = 20.18$], whereas the latter did not differ reliably ($F < 1$). Dunnett's test was used to evaluate the difference between each transfer paradigm and Control 1. The amount of retroactive inhibition was significant at the .01 level for C–D and A–C, and at the .05 level for A–Br. The difference between A–B′ and the control treatment was not reliable. The conclusions remain the same when Control 2 is used in the comparisons.

The trends over trials are considered next. For C–D the function was essentially flat; there was a substantial loss on the first test and little change on the subsequent trials. The picture is similar for A–C, although there was some evidence of a progressive

TABLE I

Mean Total Numbers of First-List Responses Recalled on Six Test Trials: Experiment I

Condition	Stringent	Lenient
C–D	37.2	39.2
A–C	37.1	39.9
A–Br	39.6	62.0
A–B′	51.1	53.1
Control 1	49.4	51.9
Control 2	55.3	56.7

decline over the first four trials, followed by a slight reversal in the trend on the two terminal tests. Thus, for the A–C as well as the C–D paradigm, the major drop in recall was observed on the first test—that is, after the study trial on the second list. By contrast, the initial drop for the A–Br paradigm was relatively small and was followed by substantial decreases on the next two tests and still another decline on the last trial. For A–B′ there was a clear increase in performance between the third and the fourth trial, and the improvement persisted on the subsequent trials. The slopes of the recall functions were significantly different from zero ($p < .01$) for A–Br and for A–B′, but not for any of the other paradigms. Analysis of the linear trends showed that there were significant differences among the interference conditions (C–D, A–C, and A–Br) and that these conditions differed reliably from A–B′ [$p < .05$ and $p < .01$ respectively].

Certain characteristics of the recall functions deserve special emphasis. First, the overall losses were no smaller for C–D than for the paradigms of associative interference (A–C and A–Br). On the early trials the difference was in the opposite direction, more so for A–Br than for A–C. This finding suggests that the initial losses represented, in large measure, a reduction in the availability of the first-list responses. The relatively high level of recall for A–Br is consistent with this assumption, since in this case the responses do not change in the transfer stage. Second, there was no clear evidence for retroactive inhibition under the A–B′ paradigm, at least if Control 1 is accepted as the appropriate baseline condition.[5] The requirement to recall both lists probably encouraged the bidirectional use of mediators (A–B–B′ and A–B′–B), with beneficial effects on both transfer learning and first-list retention. Finally, there were divergent trends over the successive test trials. The rise in recall for A–B′ is readily attributable to the development and utilization of mediational chains (A–B′–B) in the course of second-list learning. On the other hand, the decline in performance under the A–Br paradigm implies that specific first-list associations were progressively weakened as second-list learning continued. Clearly, the procedure of re-pairing makes it difficult to maintain both the first and the second response to a given stimulus. Thus, transfer learning may not progress to full mastery unless at least some of the first-list associations are extinguished or suppressed. In agreement with this interpretation, A–Br was found in a previous study to be the only paradigm for which there was a significant amount of retroactive inhibition under conditions of multiple-choice learning—that is, when correct performance reflected the integrity of specific associations and was independent of response availability.[6]

---

[5] In a previous experiment in which the study-test method was used, a significant amount of retroactive inhibition was obtained for the A–B′

An analysis was carried out to determine the relation between speed of acquisition and recall for the items in the first list. On the basis of the numbers of correct test responses during original learning, the first-list pairs were divided into three categories: items of high, medium, and low strength. Measures of recall were then determined for each of the categories. The results are not here reported in detail, but it may be stated that the differences in level of recall observed for the total lists reflected, primarily, variations in the numbers of correct responses from the categories of low and medium strength. There was relatively little forgetting for items of high strength. These conclusions hold both for the paradigms of transfer and for the two control conditions. It is clear, however, that the differences in performance were not limited to the weakest items.

The lenient scores on the successive test trials showed the same pattern as the stringent ones, except under the A–Br paradigm. Under that paradigm, the lenient scores on the initial test trial were comparable to those of the control conditions and showed

---

paradigm. See L. Postman and K. Stark, Role of response availability in transfer and interference, *J. exp. Psychol.*, 79, 1969, 168–177.

[6] Postman and Stark (see n. 5 above).

progressive increases thereafter. These gains may, of course, be attributed to the cumulative improvement in second-list performance.

Interdependence of first-list and second-list reponses. When the stimuli remain the same, the paradigm of transfer is assumed to determine the degree to which the successive responses to a particular stimulus facilitate or inhibit each other. The nature of the paradigm should, therefore, have a systematic effect on the relationship between the recall probabilities of the two responses to the same stimulus. The relevant data for the A–C, A–Br, and A–B′ paradigms are summarized in Table II. The table shows (*a*) the numbers of joint occurrences of the first and of the second response to the same stimulus, (*b*) the numbers of such occurrences to be expected if the recall probabilities were independent, and (*c*) the differences between the observed and the expected values expressed as percentages of the latter. The measures, which were based on the pooled scores of all subjects, have been grouped over Trials 1–3 and 4–6. There were pronounced differences among the paradigms on the early test trials. The relation between first-list and second-list recall was strongly positive for A–B′, in the same direction but considerably weaker for A–C, and was clearly negative for A–Br. For purposes of interpretation, it is important to recognize that the presence of a positive relation does not necessarily imply associative facilitation, since the relation may be produced by variations among stimulus terms —by responses being more easily attached to some stimuli than to others. Inferences about the operation of associative processes in the transfer stage must, therefore, be based on comparisons of the paradigms. Thus, relative to A–C there was evidence for as-

TABLE II

JOINT OCCURRENCES OF FIRST-LIST AND SECOND-LIST RESPONSES TO THE SAME STIMULUS: EXPERIMENT I

	Trials 1–3			Trials 4–6		
Paradigm	O	E	% Diff.	O	E	% Diff.
A-C	136	128.6	+ 5.8	241	226.7	+6.3
A-Br	111	126.8	−12.5	212	223.5	−5.1
A-B′	323	254.4	+27.0	419	395.4	+6.0

Note: O = Observed, E = Expected, % Diff. = 100 × (O − E / E).

sociative facilitation in A–B′ and associative interference in A–Br. The reasons for the slight positive relation observed for A–C remain uncertain.

The relative difference between observed and expected values declined on the later trials for both A–B′ and A–Br. The percentage for A–C, which was small to begin with, did not change. The possible range of the measures of interdependence was increasingly constrained as second-list learning progressed toward complete mastery.

EXPERIMENT II

The major purpose of Experiment II was to make a direct comparison of transfer and first-list recall under conditions of accretion and of substitution. A further objective was to assess the possible effect of a source of bias that may have been present in Experiment I. Specifically, Experiment I used different methods of testing in original learning and in the transfer stage: rapidly paced test trials in the former and the method of modified free recall (MMFR) in the latter. It is possible that this shift in context served to decrease the amount of retroactive inhibition and to reduce the sensitivity of the tests of recall to differences among paradigms. To check on this possibility, written tests of modified free recall were used in both stages of the present experiment. Consideration was limited to the C–D and A–C paradigms. These two treatments were chosen because they represent the conventional reference condition on the one hand and the classical paradigm of negative transfer and retroactive inhibition on the other. The previous failure to observe a difference between the two paradigms provided an additional reason for their choice.

Method

Design. There were four transfer groups, which represented the factorial combination of two paradigms (C–D and A–C) and two methods of testing (simultaneous and successive). The simultaneous method was that used in Experiment I: recall of the first as well as of the second list was required on the test trials in the transfer stage. In the successive method, the test trials in the transfer stage were limited to the second list, except for the last trial, on which recall for both lists was taken. The terminal test trials were, therefore, directly comparable under the two procedures. There were two control groups: the one for the simultaneous procedure received repeated tests on the original list, and the one for the successive procedure was tested once after the appropriate retention interval. These two conditions are designated

Control-repeated and Control-delayed respectively. Thus, the design comprised six (four experimental and two control) groups.

Lists and procedure. The lists were the same as in Experiment I. First and second lists were again fully counterbalanced. Both the first and the second lists were learned by the study-test method. On the study trials the pairs were presented at a 2-sec. rate. There were four different orders of presentation. The MMFR was used on all test trials. During first-list learning the *Ss* were allowed 1 min. to fill in the response blanks which appeared next to the stimuli on the test sheet. In the transfer stage the *S*s in the simultaneous condition were given 2 min. to write down both the first-list and the second-list responses to the appropriate stimuli. The arrangement of the test sheets was the same as was used for the corresponding transfer groups in Experiment I. In the successive conditions, recall of only the second list was required, and the tests were 1 min. in duration.

The first list was learned to a criterion of 7/12. One additional study trial was given after the criterion was attained.[7] There were seven study-test cycles in the transfer stage. During the first six cycles the test procedures differed for the simultaneous and the successive conditions as described above. On the seventh test all groups were required to recall both the first-list and the second-list responses. Thus, there was a terminal MMFR test that was directly comparable under all treatments and permitted the determination of first-list recall for the successive groups.

The Control-repeated group was given seven 2-min. tests on the original list; these were spaced so as to coincide with the corresponding tests in the transfer conditions.[8] The Control-delayed group was given a single MMFR test. The retention interval was equal to that between the end of first-list learning and the seventh test trial under the transfer treatments. All parts of that interval not used for study or test trials were filled with a digit-symbol cancellation task. It should be noted that there were periods occupied with the filler task not only for Control-delayed *S*s but also for *S*s in the successive conditions (who received 1-min. rather than 2-min. tests). The test trials for the latter conditions were spaced so as to begin at the same time as those in the simultaneous conditions. As in Experiment I, full instructions about the procedure to be followed on the test trials were given at the beginning of the transfer phase.

Subjects. There were 16 *S*s in each of the six groups. The *S*s were drawn from the same population as in Experiment I and were assigned to conditions in the same manner as before.

Results

First-list learning. The mean number of trials to the criterion of 7/12 was 3.25, with the means of individual groups ranging from 2.94 to 3.94. The mean number correct on the criterial trial was 9.71, with a range from 9.38 to 10.12. There was no significant variation among groups in speed of acquisition for either measure (*F*s of 1.12 and < 1.00 respectively). The number of trials was smaller and the terminal performance was higher than in Experiment I. However, the measures obtained in the two experiments cannot be compared directly, in view of the difference in the method of testing.

---

[7] The experimental arrangements made it necessary for the acquisition stage to end with a study trial, since there had to be time to score the last written test and to determine whether criterion had been reached.

[8] In light of the results of Experiment I it was decided to give 2-min.

Second-list learning. The acquisition curves for the second list are presented in Figure 3. The distinction between the simultaneous and the successive treatments applies only to Trials 1–6. On Trial 7 all groups were administered MMFR tests. The analysis of the transfer results is limited to the first six trials.

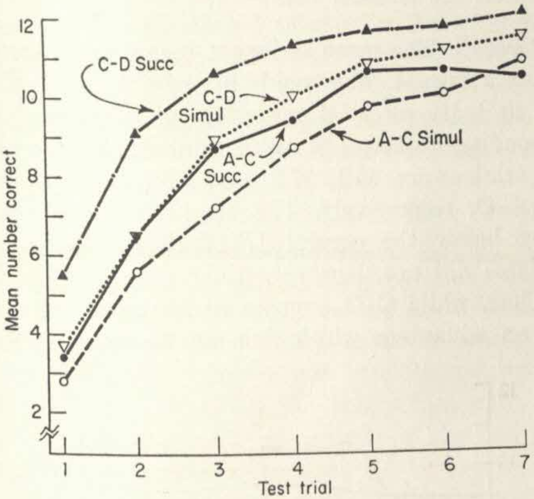


FIG. 3. Mean numbers of second-list responses given correctly on each test trial in Experiment II

Both the paradigm of transfer and the method of testing had substantial effects on speed of second-list learning. Performance was better under the successive than under the simultaneous treatment and higher for C–D than for A–C. The mean total numbers correct on Trials 1–6 in the successive condition were 59.8

---

tests to the control group in order to obtain a maximal estimate of the amount of retroactive inhibition.

for C–D and 49.2 for A–C. The corresponding values obtained in the simultaneous condition were 51.0 and 43.9. Both main effects were significant [*F*s (1, 60) of 9.40 ($p < .01$) and 5.94 ($p < .05$) for the method of testing and the paradigm of transfer respectively]. There was no reliable interaction ($F < 1$). The separation between the two conditions of testing was essentially independent of paradigm; hence, the amounts of negative transfer were closely similar. It should be noted that the difference between the two methods of testing was not great. The scores of the simultaneous group lagged behind those of the successive group by approximately one trial. Thus, the recall of the first-list associations did not seriously retard performance on the transfer list.

First-list recall. The mean stringent measures of first-list recall are shown in Figure 4. The trends in the simultaneous conditions resembled, in most essential respects, those observed earlier for the corresponding groups. The mean total numbers recalled on the seven test trials were 66.2, 57.2, and 53.4 for Control-repeated, C–D, and A–C, respectively. The combined transfer groups fell significantly below the control [$F\,(1,45) = 4.48$, $p < .05$]. The two paradigms did not, however, differ reliably from each other ($F < 1$). Thus, while C–D appears to have had some advantage over A–C, an advantage which was not present in Experiment I,

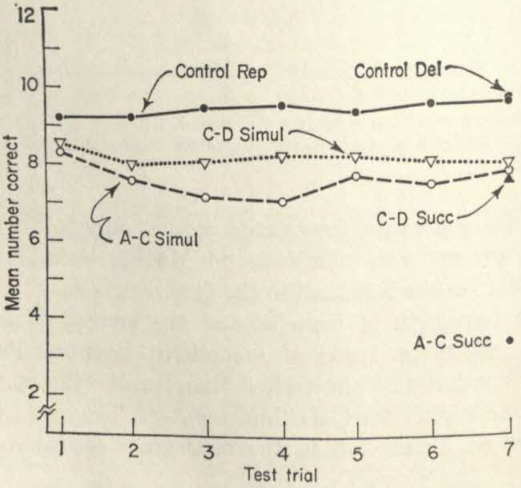


FIG. 4. Mean numbers of first-list responses recalled on each test trial in the simultaneous conditions and on the delayed test trial in the successive conditions in Experiment II

the difference was far from significant. For A–C there was again some decline on the initial trials, followed by a reversal in the trend. The function for C–D as well as for the control condition was essentially flat. Thus, the difference between A–C and C–D on the terminal trial was equal to that on the initial test.

Recall performance in the successive conditions is considered next. There was retroactive inhibition for both paradigms but, as is typically found under the conventional treatment, the loss was substantially greater for A–C than for C–D. The difference between the control and the combined transfer groups was significant, as was that between the latter [*F*s (1, 45) = 31.76 and 24.19 respectively, $p < .01$ for both].

The presence of a pronounced interaction between method of testing and paradigm is apparent. The relevant comparisons are based on the results of the terminal MMFR test. The A–C paradigm yielded a large difference between the simultaneous and the successive treatment in favor of the former; no evidence of such a difference was found in either the C–D or the control conditions. When the two transfer paradigms are compared, the interaction with the method of testing was significant at the .01 level [*F* (1, 90) = 11.09]. Performance in the relatively unfavorable successive situation necessarily set a limit on the amount of improvement that could be achieved by a shift to the simultaneous situation. Thus, a much greater change was possible and was actually observed for A–C than for C–D. The interaction may, therefore, be taken to reflect the differential sensitivity of the two paradigms to manipulations of the method of testing.

An item analysis of the stringent scores, carried out in the same manner as in Experiment I, showed that the differential effects of the method of testing were present at all levels of strength. The lenient scores were not appreciably higher than the stringent ones and yielded the same conclusions in all respects.

Second-list recall. The mean stringent measures of second-list recall in the successive condition were 12.0 for C–D and 11.0 for A–C. The corresponding values on the terminal test in the simultaneous condition were 11.4 and 10.8 (see Figure 3). The measures on the terminal test trial reflected the trends in acquisition. Considered by themselves, these differences fell short of significance.

Interdependence of first-list and second-list responses. The degree of interdependence of first-list and second-list responses to the same stimuli under the A–C paradigm was evaluated in the same manner as in Experiment I. The results are presented in Table III. In the simultaneous condition the relation observed on the first six transfer trials was again positive and somewhat stronger than in Experiment I. On the terminal MMFR test, second-list performance was so high as to preclude any substantial discrepancy between observed and expected values. For what it is worth, there was a slight indication that the relation was more likely to be positive in the simultaneous than in the successive condition. If this were the case, it would be possible to speculate that subjects are more prone to make occasional attempts at mediation in a simultaneous than in a successive situation.

TABLE III

Joint Occurrences of First-List and Second-List Responses to the Same Stimulus for the A-C Paradigm: Experiment II

	Trials 1–3			Trials 4–6		
Condition	O	E	% Diff.	O	E	% Diff.
Simultaneous	177	160.4	+10.3	307	282.2	+8.8
	Trial 7					
	O	E	% Diff.			
Simultaneous	116	112.6	+3.0			
Successive	49	50.4	−2.8			

Note: O = Observed, E = Expected, % Diff. = 100 × (O − E / E).

DISCUSSION

The results of the present experiments support three conclusions. The first of these is that the requirement to maintain and recall first-list associations during the acquisition of the second list retards the mastery of the transfer task only to a very limited extent and does not alter the characteristic pattern of differences among specific paradigms. With the exception of the paradigm of mediation (A–B′), the presence of an alternative prior association appears inevitably to retard the attachment of a new response to a stimulus. In light of the invariance of the paradigmatic relations under conditions of substitution and accretion, it seems likely that the associative interference is not a consequence of competition in the sense of the momentary dominance of one response over another. Rather, at least one important source of the retardation of second-list learning may lie in the limitations of the mechanisms that must function in the establishment of successive associations. Thus, in the case of A–C the most effective implicit associative responses that can be used to link the members of the pairs may be preempted during the acquisition of the first list. Consequently, the attachment of stable responses to the functional stimuli may be, in principle, more difficult in transfer than in original learning, and this state of affairs holds true for a simultaneous as well as for a successive procedure. If one thinks in terms of an hierarchy of effective implicit responses, the subject must sample from the lower ranks in this hierarchy as he moves from the first to the second list.

While a direct comparison between the two methods of testing was made only for the A–C and C–D paradigms (Experiment II), the results obtained in the transfer stage of Experiment I make it apparent that A–Br and A–B′ maintained their normal positions relative to the other paradigms under the simultaneous procedure. The factors that maximize associative interference under the A–Br paradigm, including the absence of consistent cues to the list membership of specific pairs, should come into play under both methods of testing. As for A–B′, it has already been suggested that the requirement to recall both lists should encourage the subject to make optimal use of the opportunities for mediational chaining. A substantial amount of positive transfer was, in fact, observed in Experiment I. The analyses of the interdependence of successive responses to the same stimuli bring out the sensitivity of the simultaneous procedure to the associative relations between lists.

The second conclusion is that retroactive interference is less under conditions of accretion than of substitution, but is not eliminated. For A–C and C–D the major retention losses occurred after the study trial on the second list, with only relatively slight declines in performance thereafter. The absence of a significant difference in the amount of retroactive inhibition between these paradigms implies that the losses reflected, primarily, a reduction in response availability. The present results do not permit any firm conclusion about the mechanism responsible for these immediate retroactive effects. It is reasonable to suppose that the study trial on the second list constitutes an occasion for the unlearning of first-list responses. During the study trial the subject necessarily focuses on the newly prescribed associations. In the interest of acquiring the new repertoire, suppression of first-list responses may occur, in spite of the subject's having been instructed that these responses will have to be recalled subsequently.

The stability over test trials of the amounts recalled in the A–C and C–D conditions presents an interesting problem of interpretation. It is tempting to infer that after the first trial, the effects of unlearning and of test-trial rehearsal balance each other. Thus, each new study trial on the second list could continue to represent a new occasion for unlearning, but there would be no decline in recall because the practice provided by the preceding test trials had increased the resistance of the first-list responses to interference. While such an exact balance of opposing influences would be surprising, it is logically possible. A plausible alternative is that the tendency to suppress first-list responses is greatly reduced, if not eliminated, as transfer learning progresses under a simultaneous procedure, whereas it continues unabated under a successive treatment.

The progressive decline over test trials under the A–Br paradigm (Experiment I) indicates that in the face of persistent heavy interference, the extinction or suppression of conflicting associations cannot be avoided indefinitely, even in the simultaneous situation. The rise in recall for A–B′ likewise reflects an increasing influence of specific associative relations over trials, in this case producing delayed facilitation rather than interference. Thus, under conditions of accretion subjects minimize or delay the suppression of conflicting associations and make progressively more effective use of the opportunities for mediation.

The third conclusion is that there are important similarities as well as differences between the characteristics of performance under conditions of substitution and of accretion. The method of testing has only a minor influence on transfer learning, but there are important effects on the maintenance of first-list associations, both with respect to the level of recall and the alignment of the paradigms of specific transfer. These findings make it apparent that, in spite of their continuity, the principles governing transfer and retroaction are far from identical. Finally, a practical conclusion that may be worthy of emphasis is that the maintenance of old associations can be achieved at relatively little cost to the efficiency of transfer performance.

THE EFFECT OF ATTENTION ON BRIGHTNESS CONTRAST AND ASSIMILATION

Leon Festinger, Stanley Coren, and Geoffrey Rivers,
New School for Social Research

Abstract. This paper reports three experiments that attempted to answer questions about the conditions under which brightness assimilation and brightness contrast are obtained. Brightness assimilation was found only under circumstances in which the gray portion of the visual display—the gray portion being compared with some other standard gray—was not the focus of attention. When attention was focused on this gray, brightness contrast was obtained. A theoretical explanation is offered in terms of the effect of attention on perceived average brightness.

The phenomenon of simultaneous brightness contrast is well known: a gray patch on a black ground appears brighter than the same gray patch on a white ground. The phenomenon is pervasive and its parameters have been thoroughly studied.[1] In addition, the facts and theories about lateral inhibition[2] seem adequate to account for brightness contrast.

At the same time there are a number of annoying facts that are not easy to reconcile with the phenomenon of contrast. These facts have also been well known for a long time. Von Bezold, in 1874, described what he called a "spreading phenomenon" that has since come to be called brightness assimilation.[3] It is, essentially, the opposite of brightness contrast, but it seems to occur only

Received for publication January 8, 1970. The study was supported by Research Grant GB-8178 from the National Science Foundation, and Research Grant MH-16327 from the National Institute of Mental Health, to Leon Festinger. The authors also wish to thank Julian Hochberg and Lloyd Kaufman for their valuable suggestions.

[1] A. L. Diamond, Foveal simultaneous brightness contrast as a function of inducing and test-field luminances, *J. exp. Psychol.*, 45, 1953, 304–314; A. L. Diamond, Foveal simultaneous contrast as a function of inducing-field area, *J. exp. Psychol.*, 50, 1955, 144–152; A. L. Diamond, Simultaneous contrast as a function of test-field area, *J. exp. Psychol.*, 64, 1962, 336–345; A. L. Diamond, Brightness of a field as a function of its area, *J. opt. Soc. Amer.*, 52, 1962, 700–706; E. G. Heinemann, Simultaneous brightness induction as a function of inducing and test field luminances, *J. exp. Psychol.*, 50, 1955, 89–96; H. Leibowitz, M. A. Mote, and W. R. Thurlow, Simultaneous contrast as a function of separation between test and inducing fields, *J. exp. Psychol.*, 46, 1953, 453–456.

[2] A. L. Diamond, A theory of depression and enhancement in the brightness response, *Psychol. Rev.*, 67, 1960, 168–199; F. Ratliff, *Mach Bands: Quantitative Studies on Neural Networks in the Retina*, 1965.

[3] W. von Bezold, *The Theory of Color and Its Relation to Art and Art-Industry*, S. R. Koehler (trans.), 1876.

in rather special circumstances. For example, if one places thin white striations on a gray background, one observes that the gray seems lighter than an identical gray with black striations on it. The phenomenon has been studied by many.[4] Among them, Helson has attempted to explain the conditions under which contrast or assimilation occurs in terms of adaptation level.[5] This explanation, however, is only partially successful. Beck, for example (see n. 4), found that Helson's theory does not account for the total range of his data.

In addition to the complexity introduced by the phenomenon of brightness assimilation, there are reports in the literature of 'cognitive' factors that affect the magnitude of brightness contrast and assimilation. Some of these reports are concerned with the effects of 'figural' qualities of the visual display. Koffka, for example, demonstrated different contrast effects on a reversible figure depending on which aspect of the visual display is seen as figure and which as ground.[6] Several investigators have also demonstrated unusual brightness-contrast effects depending on whether a gray triangle appears to be on top of a black shape or adjacent to it.[7] Other cognitive factors are related to experience. For example, Beck (see n. 4) reported that with repeated exposure to stimuli that usually produce assimilation responses, observers begin to report brightness contrast.

Coren attempted to specify, under well-controlled conditions, the effect of 'figure' on brightness contrast.[8] In one experiment he used a display that was seen as a gray rabbit on a black (or white) background. Rotated 180°, however, the display was seen as a gray space between two black (or white) faces of women. Thus, he could have the observers match the brightness of the identical gray in the identical stimulus display when that gray was figure and when it was ground. In another experiment he used stereoscopic stimuli to control which part of the display was seen as figure. A gray disc was made to stand out in front of a black (or white) ring, or the ring was made to stand out in front of the gray. Since observers perceive the part of the display that stands out in front as the figure, he could, again, have them match the brightness of the identical gray in a nearly identical stimulus display when the gray was seen as figure and when it was seen as ground. The results are rather clear. When the gray that is matched is seen as figure, there is significantly more brightness contrast than when that identical gray is seen as ground. Thus, there does seem to be a cognitive factor influencing the magnitude of brightness contrast.

---

[4] J. Beck, Contrast and assimilation in brightness judgments, *Psychon. Sci.*, 1, 1966, 342–344; R. W. Burnham, Bezold's color mixture effect, this JOURNAL, 66, 1953, 378–385; R. M. Evans, *An Introduction to Color*, 1948; H. Helson and V. L. Joy, Domains of lightness contrast and assimilation, *Psychol. Beitr.*, 6, 1962, 405–415; H. Helson and F. G. Rohles, A quantitative study of reversal of classical lightness contrast, this JOURNAL, 72, 1959, 530–538; S. M. Newhall, The reversal of simultaneous lightness contrast, *J. exp. Psychol.*, 31, 1942, 393–409; J. A. Steger, Visual lightness assimilation and contrast as a function of differential stimulation, this JOURNAL, 82, 1969, 56–72.

[5] H. Helson, Studies of anomalous contrast and assimilation, *J. opt. Soc. Amer.*, 53, 1963, 179–184; H. Helson, *Adaptation-Level Theory*, 1964.

[6] K. Koffka, *Principles of Gestalt Psychology*, 1935.

[7] W. Benary, Beobachtungen zu einen Experiment über Helligkeitskontrast, *Psychol. Forsch.*, 5, 1924, 131–142; W. T. Mikesell and M. Bentley, Configuration and brightness contrast, *J. exp. Psychol.*, 13, 1930, 1–23; J. G. Jenkins, Perceptual determinants in plane designs, *J. exp. Psychol.*, 13, 1930, 24–46.

[8] S. Coren, Brightness contrast as a function of figure-ground relations, *J. exp. Psychol.*, 80, 1969, 517–524.

To say that a cognitive factor such as the perception of figure affects simultaneous brightness contrast is interesting but not entirely satisfying. One would like to know how this cognitive factor operates, how it interacts with lateral inhibitory processes, and what the mechanisms are by means of which the magnitude of contrast is altered. A possible, relatively simple, theory suggests itself.

The visual system transmits information primarily about changes that occur, and not very much about steady retinal states. Evidence for this statement comes from both neurophysiological and psychological studies. Thus Hartline, on the basis of physiological evidence, stated that "the visual system is almost exclusively organized to detect change and motion."[9] The same conclusion was reached by others, on the basis of work with stabilized retinal images. It is well known that a stabilized image produced on the retina (so that normal eye movements no longer produce changes in stimulation) rapidly disappears.[10] In other words, if there are no changes in stimulation on the retina, visual input seems to stop. This carries the implication that in normal vision, small continual eye movements produce changes in stimulation on the retina in the neighborhood of contours—that is, the neighborhood of sharp differences in intensity. This continual change in the stimulation of retinal receptors keeps information flowing in the visual system.

---

[9] H. K. Hartline, Visual receptors and retinal interaction, *Science*, 164, 1969, 270–278, at p. 275.

[10] R. M. Pritchard, W. Heron, and D. O. Hebb, Visual perception approached by the method of stabilized images, *Canad. J. Psychol.*, 14, 1960, 67–77; L. A. Riggs, F. Ratliff, J. C. Cornsweet, and T. N. Cornsweet, The disappearance of steadily fixated visual test objects, *J. opt. Soc. Amer.*, 43, 1953, 495–501.

But if the visual system does not in fact transmit much information about steady retinal states, the attempt to explain normal visual experience presents some problems. For example, what happens if an observer looks at a large black square on a white background, maintaining reasonable fixation in the center of the square? Presumably there is considerable information input from the contour but little or none from the center of the uniform black square. How, then, does the observer see a uniform black square? It must be that the central nervous system, in the absence of reliable input from some area, assumes uniformity between contours.

A convincing demonstration of this process was provided by Krauskopf.[11] The observer in this experiment is presented with a stabilized disc surrounded by a nonstabilized colored annulus. After a few seconds the stabilized disc fades and disappears. But what does the observer see then? He does not, of course, see an empty hole in a colored annulus. He simply sees a uniformly colored circle in his visual field. The same result has also been reported by Yarbus and Gerrits.[12] Krauskopf (see n. 11) summarized the theoretical conclusions to be drawn as follows:

> It would seem that information indicating the existence of contours between regions of the visual field determine how the regions themselves are perceived. Under normal fixation conditions, responses generated by the movement of the disk-annulus border over the receptors signal the existence of a change in stimulation between the disk and annulus. Under prolonged stabilized viewing, such information is absent and the whole field is seen in the color of the annulus since there only is information concerning the change in stimulation between the surround and the annulus. (p. 743)

One might be tempted to maintain, on the basis of this kind of evidence, that no information at all is transmitted about steady states on the retina. This is not a plausible assertion, however, considering other neurophysiological evidence. Microelectrode recordings in the optic tract and in the lateral geniculate show that the firing rates for steady states are directly related to the intensity of stimulation on the retina.[13] Nevertheless, these differences are small compared to the transient responses that signal the magnitude of change. We can at least maintain, therefore, that information about steady states is relatively poor and unreliable.

---

[11] J. Krauskopf, Effect of retinal image stabilization on the appearance of heterochromatic targets, *J. opt. Soc. Amer.*, 53, 1963, 741–743.

[12] A. L. Yarbus, *Eye Movements and Vision*, B. Haigh (trans.), 1967; H. J. M. Gerrits, Observations with stabilized retinal images and their neural correlates, doctoral dissertation, Catholic University of Nijmegen, 1967.

If the visual system does not transmit much reliable information about steady states, but only about changes that occur, then what determines the perception of absolute brightness levels? We would like to propose that the visual system takes a crude average of the relatively unreliable input about steady states across the entire visual field to establish an absolute brightness level. We would further like to suggest that areas of the visual field with 'figural' characteristics are overweighted in the computation of this crude average. Differences in brightness—that is, the changes in stimulation produced in the neighborhood of contours by continual eye movements—are then superimposed on this weighted average of brightness.

The proposal that absolute brightness level is derived from a weighted average over the entire visual field is not a new idea. For example, in the attempt to explain the phenomena of brightness and color constancy, investigators such as Katz and Bühler proposed that the observer directly perceives the absolute level of illumination, this perception being derived from the entire visual field.[14] More recently, Helson, addressing himself to the same problem, stated that "background reflectance, by virtue of the large area of background and because background furnishes the border for all samples in the field, is the most important single factor in the visual field determining *adaptation reflectance which is to be regarded as a weighted mean reflectance of all parts of the visual scene*" (italics ours).[15] In general, of course, Helson's concept of adaptation level is similar in nature to our own proposal.

If such an overweighting of 'figure' occurs, the average brightness of a display in which the figure is brighter than the ground would be raised somewhat. The average brightness would be lowered on displays in which the figure was darker than the ground. Assuming that the information about differences in brightness is symmetrically superimposed on this weighted average of brightness, this would result in displacement of the brightness of all parts of the visual display. We can thus deduce the following effects. If a gray figure on a white ground is compared to an identical gray figure on a black ground, the perceived brightness of the former gray would be less than that of the latter gray. Thus, for figures one would observe brightness contrast. On the other hand, if one compares a gray ground with a white figure on it to an identical gray ground that has a black figure on it, the former gray would be perceived as brighter than the latter gray. Thus, for ground one would observe brightness assimilation. The combined effects of this process and the processes of lateral inhibition that push toward contrast might be expected to produce stronger effects for brightness contrast than for brightness assimilation.

---

[13] O. Creutzfeldt, J. M. Fuster, A. Herz, and M. Straschill, Some problems of information transmission in the visual system, in *Brain and Conscious Experience*, J. C. Eccles (ed.), 1966.

[14] D. Katz, *The World of Color*, 1935; K. Bühler, *Handbuch de Psychologie*, 1922.

[15] H. Helson, Some factors and implications of color constancy, *J. opt. Soc. Amer.*, 33, 1943, 555–567, at p. 562.

The preceding analysis suggests that 'figure' contrasts from 'ground' and that 'ground' assimilates to 'figure.' However, Coren reported no instances of brightness assimilation in his data (see n. 8). He found brightness contrast for ground as well as for figure. His data showed only that there is more contrast when the test gray is figure.

Let us examine what is meant by 'figure' and by 'ground,' and consider why there might be a difference in the weighting given to these different parts of the visual field. We generally denote as figure that part of the visual field which captures the attention of the observer. This is the part of the display that he 'looks at,' that he examines, to which he is prepared to respond. The rest is background, to which he 'pays less attention.' Let us propose that it is the act of attention that produces the overweighing in the absolute brightness averaging and not the quality of 'figure' per se. Our theoretical suggestion then can be revised as follows: That part of the visual field which captures attention shows the phenomenon of *brightness contrast;* those parts of the visual field which are *not* attended to are likely to show *brightness assimilation.*

If this is a correct formulation, we can then offer a tentative explanation of why Coren found no brightness assimilation in his study (see n. 8). If the observer is asked to match the brightness of a test gray with a variable gray, then regardless of whether that test gray is figure or ground, regardless of whether it would normally capture his attention or not, he is forced to pay some attention to the gray he is required to match. Thus, since the observer is always paying some attention to the gray, Coren obtained brightness contrast in all conditions. The distribution of attention over the display, however, is not the same when the gray is figure and when the gray is ground. This results in a difference in the magnitude of brightness contrast among the different conditions.

If this relatively simple hypothesis can integrate and explain brightness contrast, brightness assimilation, and the effect of some 'cognitive' factors on these phenomena, then it must, of course, be able to deal with the known circumstances under which brightness assimilation is normally obtained. These known circumstances should turn out to be instances in which the observer does not carefully attend to the test gray he is judging. Therefore, let us consider in detail those displays that normally produce brightness assimilation.

The reader will recall that the kind of display which produces assimilation responses is one in which there are thin white (or black) striations on a gray ground. This has been shown both by Helson and Rohles and by Helson and Joy (see n. 4). Their data show that when the black or white stripes were thinner than the interspaced gray stripes, brightness assimilation was obtained. When, however, the gray stripes are thinner than the black or white stripes, brightness contrast is produced. Thin lines on a display are more likely to capture the attention and to be seen as 'figure,' thus producing the contrast or assimilation results that are found. Several investigators studied this question systematically and found, indeed, that the thinner portions of a reversible stimulus are more likely to be seen as figure than its broader portions.[16]

We come to the conclusion, then, that displays which normally produce brightness assimilation are ones in which the test gray is seen as background. However, in explaining Coren's data we conjectured that asking the subject to match a test gray must force his attention onto that gray to some extent. How do displays that consistently produce brightness assimilation avoid this problem?

---

[16] C. H. Graham, Area, color, and brightness difference in a reversible configuration, *J. gen. Psychol.*, 2, 1929, 470–483; H. Goldhamer, The influence of area, position, and brightness in the visual perception of a reversible configuration, this JOURNAL, 46, 1934, 189–206; T. Oyama, Figure-ground dominance as a function of sector angle, brightness, hue, and orientation, *J. exp. Psychol.*, 60, 1960, 299–305.

Let us examine, in detail, the procedures used in experiments that report brightness assimilation. These experiments typically present two gray rectangles side by side, one bearing thin white lines and the other bearing thin black lines. The observer is not asked to do any matching but is simply asked to report which gray looks lighter. In other words, the response asked for does not require very careful attention to the gray. In addition, the stimulus-exposure times are always kept brief, typically about three seconds. In short, when the test gray is ground and the response required does not force attention to the test gray, and when the presentation time is brief enough so that the figure captures the attention effectively for that period of time, then one obtains brightness assimilation.

Our speculations are, of course, amenable to experimental testing. If the task facing the observer is one that forces attention to the test gray, then figures that normally produce assimilation should show contrast.

EXPERIMENT I

This experiment was designed to answer the question about the effect of the method of measurement on whether one observes brightness contrast or brightness assimilation. The methods of paired comparison with brief exposures and of brightness matching were employed, both with stimuli that have been used in assimilation studies and with stimuli that typically produce brightness contrast.

Method

Stimuli. The *'assimilation stimuli'* were 10-cm. squares of gray paper (35% reflectance) with regularly spaced black (2.4% reflectance) or white (82% reflectance) vertical lines that were 6 mm. wide. The intervening gray stripes were 12 mm. wide. The *'contrast stimuli'* were the same size and consisted of a gray vertical bar, 38 mm. wide, in the center, flanked by two black or two white bars, each being 31 mm. in width. Five *practice stimuli* were also used. Three of these were uniform grays: one 19% reflectance, another 35% reflectance, and the third 50% reflectance. The two other practice stimuli contained either a white or a black 38-mm. square in the center of a 35%-reflectance gray.

Apparatus. The observer viewed the stimuli through a 23 cm. × 38 cm. rectangular aperture equipped with a manual shutter. The stimuli were mounted 85 cm. behind the aperture on a black (2% reflectance) background.

For paired-comparison judgments two stimuli were displayed simultaneously 5 cm. apart. To obtain brightness matches, one stimulus was presented at a time. The observer, by turning a handle mounted on his right, could vary the size of the black and white sectors on a spinning Gerbrands differential rotor. Readings were taken in degrees of white from the rotor shaft and later converted to percent reflectance.

The light incident on the stimuli was provided by a ring of ten 15-w. tungsten light bulbs mounted out of view behind the rectangular aperture. A filtered d.c. power source was used in order to eliminate stroboscopic effects on the rotor face. The lights produced a uniform flux of 50.3 ftc. at the plane of the stimuli.

Subjects. Ten paid volunteers were recruited on the campus of Stanford University. All had 20/20 visual acuity, normal or corrected.

Procedure. Each *S* made both paired-comparison judgments and direct brightness matchings of the stimuli. Half of the *S*s did the paired comparisons first, and the other half did the brightness matchings first. For the brightness-matching situation, one stimulus was presented at a time and *S* set the rotor so that it matched the gray of the stimulus. Two matches were made, one starting with the rotor face obviously darker than the gray and another starting with the rotor face obviously lighter. The order of stimulus presentation follows. Each *S* first matched the uniform practice grays of 19% and 50% reflectance in mixed order. Matches were then made for the 35%-reflectance uniform gray, the same gray as on the test stimuli. The four test stimuli—two assimilation stimuli and two contrast stimuli—were then presented in mixed order.

To obtain paired comparisons, two stimuli—one with white and one with black—were simultaneously presented for 3 sec. and *S* was asked to state on which side, left or right, the gray was lighter. He was told to guess if uncertain. The two assimilation stimuli were presented side by side four times, and the two contrast stimuli were presented together four times. Two other pairs—one consisting of the black square on gray and the white square on gray, another of the 19%- and 50%-reflectance uniform grays—were each presented twice. The order of presentation of these stimulus pairs was random. Which stimulus in the pair appeared on the left or the right was balanced.

Results

With judgments made using the paired-comparison procedure and a 3-sec. exposure, we would expect the 'assimilation stimuli' to produce assimilation responses (the gray with white to be judged lighter) and the 'contrast stimuli' to produce contrast responses (the gray with black to be judged lighter). This is what was found, as Figure 1 shows. Seven subjects gave contrast responses on all four presentations of the contrast stimulus pair. Two gave three contrast responses and one assimilation response, and only one subject gave one contrast and three assimilation re-

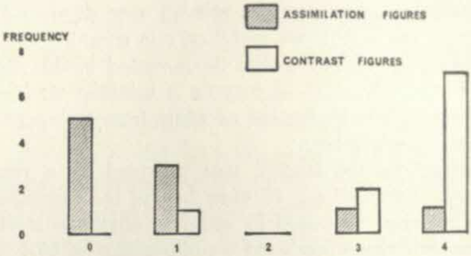


FIG. 1. Distribution of number of contrast responses out of four stimulus presentations for three-second exposure; paired-comparison technique

sponses (significantly different from chance, $p < .01$, Kolmogorov-Smirnov test).[17] The picture is quite different for the assimilation stimulus pair. Five subjects gave assimilation responses on each of the four presentations; three did so on three of the four presentations. Only two subjects gave more contrast than assimilation responses (significantly different from chance, $p < .05$, Kolmogorov-Smirnov test). The two distributions were, of course, significantly different from each other ($p < .01$).

These results simply replicate what has been reported in the literature.[18] With relatively thin black or white lines on a gray background, and with brief exposure in a paired comparison, one obtains brightness assimilation. The important point comes in comparing these results with the results obtained on the same stimulus figures using brightness matching. And as Table I shows, these results were quite different from those obtained by paired comparison. Here there was no longer any difference in the results produced by the 'assimilation' and by the 'contrast stimuli.' Both stimulus types produced brightness contrast. The gray with the white inducer was seen as significantly darker than the gray with the black inducer for both the contrast ($p < .01$, $t = 5.91$) and the assimilation stimuli ($p < .01$, $t = 7.60$). Using this procedure for measurement, brightness assimilation did not occur.

Discussion

The data are consistent with the idea that the difficulty in obtaining brightness assimilation, even when the test gray is background, lies in the measurement procedure that forces the observer

---

[17] S. Siegel, *Nonparametric Statistics for the Behavioral Sciences*, 1956.
[18] See Helson (n. 5 above).

TABLE I

MEAN MATCH SETTINGS ON DIFFERENTIAL ROTOR, IN PERCENT REFLECTANCE, FOR TEST GRAY IN CONTRAST AND ASSIMILATION CONFIGURATION: EXPERIMENT I

	Configuration		
	Contrast	Assimilation	Base gray
White inducer	20.2	20.8	35.1
Black inducer	38.2	34.5	

to pay attention to that gray. This could well be the reason that Coren did not obtain anything but brightness contrast (see n. 8). Thus, it is still possible to maintain the hypothesis that the part of the visual field to which the observer pays attention shows contrast effects while the part to which he does not attend shows brightness assimilation.

Presumably, the reason that we did obtain evidence of brightness assimilation with the proper stimuli on a 3-sec. exposure is that the thin black or white lines captured the observer's attention and the exposure time was too short for him to redirect his attention to the gray to which he was supposed to be responding. If this is a correct interpretation, then one might expect that, even using a paired-comparison procedure, the assimilation responses would tend to disappear if the presentation times for the stimuli were longer. With a longer presentation time, the observer would be able to shift his attention to the gray, and if this happened, the measurements would show brightness contrast. Experiment II was designed to investigate this question.

EXPERIMENT II

Method

The stimuli, the apparatus, and the general procedure were all similar to the paired-comparison portion of Experiment I. In this experiment the illumination incident on the stimulus plane was 30 ftc. In addition to a condition under which the pairs of *'assimilation'* and *'contrast stimuli'* were exposed for *3 sec.*, another condition was run in which the same stimulus pairs were exposed for *10 sec*. In this latter condition the *S*s were told to pay careful attention to the gray on the stimuli and were instructed not to respond until the shutter was closed at the end of 10 sec.

Subjects. Twenty *Ss* with 20/20 normal or corrected vision were recruited from the New School for Social Research. Ten of them, randomly assigned, were in the 3-sec. exposure condition and the other ten were in the 10-sec. exposure condition.

Results

The results, in terms of the number of contrast responses obtained in the four presentations of the stimulus pairs, are presented in Figure 2. It is clear that for the 3-sec. exposure time, the results closely replicate the findings from Experiment I. All of the statistical comparisons, analyzed as in Experiment I, were also similarly significant. Again, we obtained assimilation responses from the 'assimilation stimuli' ($p < .05$) and contrast responses from the 'contrast stimuli' ($p < .01$).

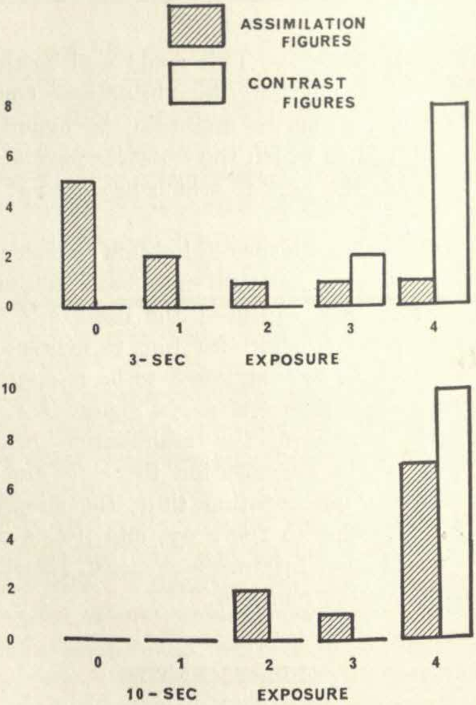


FIG. 2. Distribution of number of contrast responses out of four stimulus presentations as a function of exposure duration; paired-comparison technique

The results for the 10-sec. exposure condition are very different.

When the subject was encouraged to pay attention to the gray, and when enough time was provided for the subject to attend to the gray, both types of stimuli yielded primarily contrast responses. The difference in responses to the assimilation stimulus pair between the two experimental conditions was significant ($p < .01, t = 5.55$).

Discussion

Experiments I and II make the same single point. If one provides conditions that direct the observer's attention to the gray in those stimulus configurations which presumably produce brightness assimilation, one then observes only brightness contrast. Brightness assimilation seems to occur if, and only if, the observer's attention is caught and held by that part of the visual display which is *not* being judged. In the case of these so-called assimilation stimuli, the parts of the display that catch and hold the attention for brief periods are the thin black or white striations.

If our hypothesis of the effect of attention on whether one obtains contrast or assimilation is correct, however, the crucial aspect is not the existence of thin black or white striations. The crucial aspect is, rather, in capturing and holding the attention of the observer so that the test gray is not attended to even though it is the part of the display that must be judged. One should be able to devise other stimulus configurations that, at least for short periods of time, also attract and hold the attention of the observer. If our explanation is correct, these should also produce assimilation responses. And since moving objects in the visual field tend strongly to capture the attention of an observer, we should be able to use this property of movement to hold attention and thereby to produce brightness-assimilation responses for stimulus configurations that would normally show brightness contrast. Experiment III was designed to examine this question.

EXPERIMENT III

This experiment compared the responses to four different stimulus-pattern and stimulus-presentation conditions. For some stimuli, the figure was gray and the background was black or white. Since the figure captures the attention and the figure is the area to be judged by the observer, only contrast responses should be obtained here, whether the figure is stationary or moving. For other stimuli, the figure was black or white and the background was gray, the gray being the part of the display to be judged. When these latter stimuli are exposed for 10 sec., there is sufficient time for attention to be shifted to the gray, and hence we would expect primarily contrast responses. When they are exposed for only 3 sec., there is less certainty that the observer has time to attend to the gray, and hence we would expect sometimes to obtain brightness contrast and sometimes brightness assimilation. And if, with these same stimuli presented for 3 sec., the black or white figure is made to move continually, this should help to capture and hold attention to the figure. Under these circumstances we would expect the gray to show primarily brightness assimilation.

Method

Stimuli. Each stimulus subtended a visual angle of 9°10′ in width and 6° 28′ in height. It was divided into two equal parts by a vertical black line 10′ wide. Each stimulus had a pattern of gray and white on one side and an identical pattern with black instead of white on the other side. The gray was always 31% reflectance. Depending on the material used for constructing the stimuli, the blacks varied from 1.5% to 2.4% reflectance and the whites from 86% to 88% reflectance.

Four *'contrast stimuli'* were prepared, using a gray figure on black and on white backgrounds. Two of these, a star and an H were *stationary*. For the other two, a circle moved from right to left and back, or a square moved up and down. The *moving* stimuli were made by preparing two or three stimuli in each of which the figure was in a different position so that when presented in succession at proper temporal intervals, apparent movement was seen.

Twelve sets of *'potential-assimilation stimuli'* were prepared. These contained on one side a black, and on the other side an identical white, figure on a gray background. Each of these sets could be presented as a *moving* figure, or one of the set could be presented as a *stationary* figure. Thus, for example, one set showed a 0 changing into a 3, which then changed into an 8 and then changed back again. For a comparable stationary figure, only the 8 was presented. In another set an arrowhead could be made to flip back and forth from left to right. For a stationary figure an arrowhead pointing in just one direction was used. Two *practice stimuli* were also used. These had physically unequal grays on the two sides.

Apparatus. The stimuli were presented in a three-channel tachistoscope (Scientific Prototype Model GB320). The light incident on the plane of the stimulus was adjusted to 35.5 ftc. for each channel.

Subjects. Twenty-six paid volunteers were recruited from the New School for Social Research. All had 20/25 or better visual acuity, normal or corrected. Two of these *Ss* gave incorrect responses on the practice stimuli with unequal grays and were discarded without further collection of data. This left 24 *Ss* in the experiment proper.

Procedure. Each *S* was shown the four contrast stimuli and the twelve potential-assimilation stimuli. For any one *S*, four of the potential-assimilation stimuli were presented stationary for *10 sec.*, four were stationary for *3 sec.*, and four were moving for *3 sec.* The presentation mode was balanced so that over all *Ss*, each stimulus was in each mode equally often. The contrast stimuli were all presented for 3 sec. To minimize any possible order effects, the 16 stimuli were presented in blocks of four, each block containing one contrast stimulus, one 3-sec., one 10-sec., and one moving potential-assimilation stimulus. The order of presentation within each block was also balanced.

Each *S* was instructed not to respond until the termination of each stimulus presentation. He was then to describe the stimuli and to indicate on which side of the stimulus the gray was lighter. He was asked to guess if uncertain. The instruction to describe the figures was intended to heighten the likelihood that *S* would pay some attention to the figure on each stimulus.

Results

Each subject made four judgments in each stimulus-presentation mode. A score from zero through four was given to each subject for each mode according to the number of contrast responses. If the subject did not give a contrast response, it was, of course, an assimilation response. Table II presents the means and standard deviations of this measure for the four presentation modes.

The 'contrast stimuli,' whether moving or stationary, yielded contrast responses almost exclusively. Twenty-two of the subjects gave four contrast responses; the other two subjects gave three contrast responses out of a possible four. Thus, when the gray to be judged was also the figure, so that all attention was centered on it, unequivocal brightness contrast was obtained.

TABLE II

MEAN NUMBER OF CONTRAST RESPONSES OUT OF FOUR STIMULUS PRESENTATIONS: EXPERIMENT III

	Contrast figures	Assimilation figures		
	3-sec.	10-sec. stationary	3-sec. stationary	3-sec. moving
Mean number of contrast responses (*SD*)	3.92 (.28)	2.63 (1.21)	1.96 (.81)	1.17 (.82)

The 'potential-assimilation stimuli,' when presented for 10 sec. as stationary patterns, still yielded primarily contrast responses, although considerably fewer than did the contrast stimuli. The difference between these two was significant ($p < .01$ using a sign test). Eighteen subjects showed fewer contrast responses, five showed the same number, and only one showed more, on the 10-sec. stationary exposure than on the contrast stimuli. Contrast responses were obtained significantly more often than chance, however ($p <$ .05). Fifteen subjects gave three or four contrast responses, while only four of them gave zero or one. The difference between the contrast stimuli and the 10-sec. stationary potential-assimilation stimuli was as expected, although not compelling from a theoretical point of view. After all, the amount of black or white that presumably produces the contrast was markedly different between these two sets of stimuli.

The comparison between the 10-sec. and the 3-sec. stationary conditions is more relevant theoretically. Here the subjects were comparing identical stimuli, the only difference being the duration of presentation. When these stimuli were presented for only 3 sec., significantly fewer contrast responses were obtained ($p < .01$). Seventeen subjects showed fewer contrast responses, four the same number, and only three give more contrast responses on the 3- than on the 10-sec. exposure. This again was in line with expectation. If the exposure time is so brief as to interfere with the transfer of attention from the figure to the gray, we would expect fewer reports of contrast. Again, however, the comparison was not compelling since we did not obtain, on the 3-sec. exposure, significantly more assimilation than chance would allow. The average was almost exactly 2.0, and one might argue that in a brief exposure, with little opportunity for examination, observers simply guessed, thus yielding chance results. This seems implausible since the contrast stimuli were also only presented for 3 sec., and the responses were clearly contrast responses. Nevertheless, it is a possibility.

The critical comparison lies between the 3-sec. stationary condition and the 3-sec. moving condition. This difference was highly significant ($p < .01$). Seventeen subjects gave fewer contrast responses, six the same number, and only one gave more on these moving stimuli than on the stationary ones. Furthermore, with the moving stimuli, significant evidence of brightness assimilation was obtained ($p < .01$). Only two subjects gave three or four contrast responses. Eighteen subjects gave zero or one contrast response. In other words, with this brief, 3-sec. exposure, and with a stimulus where the figure so moved as to capture the attention of the observer, brightness assimilation with stimulus patterns that are not normally considered assimilation stimuli was obtained.

Discussion and Conclusions

The data from all three experiments are consistent with the theoretical explanation that we advanced. It is necessary, however, to look at other possible interpretations. The main alternative explanation that suggests itself concerns possible differential eye-movement patterns between different conditions. After all, eye movements and fixation points were not controlled in any of the experiments; the observers were free to move their eyes at will. It is quite plausible to suppose that an observer's eyes fixate differently depending upon what attracts his attention and depending upon the task. There are two separate ways in which such eye-movement differences could affect measurements of brightness contrast. First of all, there could be differences in magnitude of contrast depending upon the position of the image on the retina. Secondly, if the eye movements are different, the sequence of successive brightness contrasts could be different. Since our measures undoubtedly reflect a combination of simultaneous and successive brightness contrast, this might explain our data. Let us consider each of these possibilities.

As to the first possibility, let us begin by saying that we have found contrast effects if the observer pays attention to the test gray, and assimilation effects if the observer's attention is held away from the test gray. If we attempted to explain this result in terms of eye fixation and consequent different regions of the retina on which the image falls, we could restate it as follows. If the test gray falls in or near the fovea, one obtains contrast, and if the test gray falls on the retinal periphery, one obtains assimilation. The results would be adequately explained if one found that brightness-contrast effects were strongest at or near the fovea.

Unfortunately, we have not been able to find any good data in the literature that bears on this question. What little we have been able to find seems to indicate the opposite. Tschermak, considering the Schumann grid effect, came to the conclusion that contrast is stronger on the periphery of the retina than on or near the fovea.[19] More recently, Alpern measured the magnitude of metacontrast with the retinal position of the test field varying from foveal to 4.25° off the fovea and found steady increases in the magnitude of metacontrast as the test patch moves away from the fovea into the periphery.[20] This perhaps supports the conclusion reached by Tschermak, but it seems highly unlikely that the mechanisms involved in metacontrast are the same as those in simultaneous brightness contrast. From existing data, it is not possible to reject completely the explanations based upon difference in fixation patterns between the various conditions. On the other hand, indirect evidence makes it seem implausible.

---

[19] A. Tschermak, Über kontrast und irradiation, *Ergebn. Physiol.*, 2, 1903, 726-798.

As to the second possibility, let us ask whether the data can be explained in terms of successive contrast effects. If an observer fixates first, say, a white figure and then fixates a gray ground, successive contrast would occur and, presumably, produce a more marked contrast response. Perhaps in a 10-sec. exposure this occurs more frequently than in a 3-sec. exposure, since there is more time for such eye movements. It is, however, difficult to see how this process could produce assimilation responses. In addition, it does not seem that this process adequately explains the large number of assimilation responses obtained with moving stimuli, for when the white figure moves, there is a period of time in which that part of the retina previously stimulated by white is stimulated by gray. This, presumably, should add to the contrast effect rather than produce assimilation. It does not seem plausible to the authors that all of the results reported above can be explained in terms of different eye-movement patterns.

In order to explain the known facts about the perception of brightness, it is necessary to begin to formulate a theory about the processing of information in the visual system. What information does and does not get transmitted? What is done with the information that is transmitted? We have brought together a number of statements to form a partial theory about visual information processing. Few of these statements are new, but bringing them together seems to help explain the seemingly contradictory phenomena of brightness contrast and brightness assimilation. This theory can be summarized as follows:

1. The visual system transmits information primarily about changes that occur on the retina and transmits little information about steady states.

---

[20] M. Alpern, Metacontrast, *J. opt. Soc. Amer.*, 43, 1953, 648–657.

2. Because of continual eye movements, changes in stimulation on the retina occur in the neighborhood of contours—in the neighborhood of sharp intensity (or wavelength) differences.
3. The visual system interpolates between contours and assumes uniformity of stimulation in areas from which little or no information arrives.
4. Absolute brightness levels are arrived at by an averaging over the entire visual field.
5. Those areas to which the observer pays attention are overweighted in arriving at this averaged, absolute, brightness level.
6. Information about changes (magnitude of difference on two sides of a contour) is symmetrically superimposed on this absolute brightness level.

We believe that this model can account for much of the data on brightness contrast and brightness assimilation. Furthermore, we have shown that by controlling exposure conditions in accordance with implications from this model, one can produce contrast with stimuli that normally yield assimilation and assimilation with stimuli that normally produce contrast.

IMPLICIT SPEECH AS AN ENCODING MECHANISM IN VISUAL PERCEPTION

ROBERT L. COLEGATE and CHARLES W. ERIKSEN,
University of Illinois at Urbana-Champaign

Abstract. The present experiment investigated the possibility that information encoding in visual perception consists of implicit speech. Since one-syllable words can be spoken, implicitly as well as explicitly, much more rapidly than three-syllable words, it was anticipated that subjects using one-syllable responses could encode more information from a rapidly decaying icon than those using three-syllable responses. Three or six nonsense forms were simultaneously presented in a tachistoscopic exposure. One group of subjects had previously learned one-syllable names for the nonsense forms; and another group, three-syllable names. At temporal intervals after termination of the display that were too long for iconic storage to persist, a probe designated the location of one of the elements in the display and the subject was asked to report what nonsense form had been in this location. The one-syllable response group was found to have encoded significantly more information from the display. However the superiority of the one-syllable response group was not as large as would be expected from the differences in duration of implicitly speaking a one- as opposed to a three-syllable response.

Although differing in detail, most current models of information processing in visual perception distinguish several substages. There are the sensory channels through which the information is conveyed to higher centers, where it persists for a brief duration as an image, or icon, before decaying rapidly. During that brief duration, a scanning, or noting, process encodes the information into a short-term memory system. Recent research has provided information on characteristics of the icon as well as on short-term memory storage.[1] However the nature of the noting or encoding process is still not clearly understood. Neisser has speculated that the process of encoding may be verbal,[2] and although it is a rather crude

Received for publication January 6, 1970. The investigation reported was supported by United States Public Health Service Research Grant MH-1206 and Research Career Award K6-MH-22014.

[1] E. Averbach and A. S. Coriell, Short-term memory in vision, *Bell System tech. J.*, 40, 1961, 309–328; C. W. Eriksen and J. F. Collins, Temporal course of selective attention, *J. exp. Psychol.*, 80, 1969, 254–261; G. Sperling, The information available in brief visual presentations, *Psychol. Monogr.*, 74, 1960 (No. 11, Whole No. 498).

[2] U. Neisser, *Cognitive Psychology*, 1967.

analogy, the notion that some part of the encoding process proceeds as though one were saying to oneself the name of the object seems to have some introspective validity. Objective support for this possibility is found in experiments demonstrating that the most likely confusions in short-term memory for visually presented items are between those items whose names sound alike.[3]

The present experiment explored the possibility that the encoding process from iconic storage to short-term memory consists of implicit naming responses, or vocalizations. Landauer has shown that implicit speech occurs at about the same rate as overt speech.[4] Further, the more syllables a word has, the longer it takes to say the word, whether the speaking is implicit or overt. Since iconic decay is quite rapid, generally being complete within 500 msec.,[5] monosyllabic names for stimuli should permit a greater number of stimuli to be encoded before the icon has decayed beyond the retrievable stage than multisyllabic names if, indeed, the encoding does consist of implicit speech.

Accordingly, three or six nonsense forms were simultaneously presented in a tachistoscopic exposure. One group of subjects had previously learned one-syllable names for the nonsense forms, and another group, three-syllable names. At temporal intervals after termination of the display that were too long for iconic storage to persist, a probe, or indicator, designated the location of one of the elements in the display and the subject was asked to report what nonsense form had been in this location. It was assumed that the more nonsense forms the subject had encoded from the display, the higher would be the probability that the indicator point to one of the nonsense forms that he had encoded. It was further assumed, and this was the primary interest of the experiment, that the subjects in the one-syllable response group would have been able to encode a larger number of nonsense forms from the displays before the icon decayed beyond the retrievable point than would those subjects in the three-syllable response group.

---

[3] A. D. Baddeley, Short-term memory for word sequences as a function of acoustic, semantic, and formal similarity, *Quart. J. exp. Psychol.*, 18, 1966, 362–365; R. Conrad, An association between memory errors and errors due to acoustic masking of speech, *Nature*, 193, 1962, 1314–1315; W. A. Wickelgren, Acoustic similarity and intrusion errors in short-term memory, *J. exp. Psychol.*, 70, 1965, 102–108.

[4] T. K. Landauer, Rate of implicit speech, *Percept. mot. Skills*, 15, 1962, 646.

[5] Eriksen and Collins (see n. 1 above).

METHOD

Subjects. Five men and seven women, students at the University of Illinois, served as paid *S*s. All had normal or corrected-to-normal vision.

Apparatus and stimuli. Stimuli were presented with a Scientific Prototype Model GA three-field tachistoscope, which had been adapted to use Sylvania F4 T5/CWX fluorescent lamps. The adaptation field was kept dark except for a transilluminated fixation cross subtending .5° of visual angle at a luminance of .5 mL. The experimental stimuli were eight nonsense forms (six- to eight-sided), constructed according to the method of Attneave.[6] Each form subtended .5° of visual angle on its maximum dimension, and they were arranged around an imaginary circle of 2½° diameter, centered on a fixation point. These forms were presented to *S* in either three- or six-form displays. For the six-form displays, the nonsense forms were located at the positions corresponding to 2, 4, 6, 8, 10, and 12 o'clock, and at the 4, 8, and 12 o'clock positions for the three-form displays. For each display size a set of eight cards was prepared by photographing black paper cutouts of the nonsense forms on a white background. In constructing these three- and six-form displays, the forms were presented within the constraints that no form occur more than once on a display, that each form occur equally often in each set, and that each form occupy a given position equally often.

The third field of the tachistoscope was used to present the probe, which consisted of a bar designating one of the positions on the display. It was presented transilluminated in an otherwise dark field at a luminance of .5 mL. It was located outside of the imaginary circle on which the forms were located and was positioned so as to constitute an extension of the radius of the circle through the designated form's position. Its closest point was .33° of angle from the designated form. By transilluminating the probe, luminance summation–contrast reduction effects for the nonsense forms and the display were eliminated.[7]

The one- and three-syllable names for the eight nonsense forms were *fan, boom, hem, con, cab, rat, dig, sup,* and *fantasy, boomerang, hemisphere, confiscate, cabinet, ratify, dignity,* and *supplement.* These words had previously been used by Eriksen, Pollack, and Montague,[8] and met the following criteria: (*a*) each of the one-syllable words was the first syllable of the three-syllable word; (*b*) each of the three-syllable words was accented on the first syllable; (*c*) all the three-syllable words were words in which all three syllables are enunciated in common usage; and (*d*) one- and three-syllable word pairs were matched for Thorndike-Lorge frequencies.

Procedure. Before beginning the experiment proper each *S* was given three practice sessions on a related tachistoscopic task using six nonsense forms (six- to nine-sided) different from those used in the experiment proper. The *S* was presented a single form at a brief duration and then 1 sec. later was presented a four-form display. He was required to report which of the four forms was the one that had been previously presented. Besides providing familiarization with tachistoscopic tasks, these practice sessions were used to determine an exposure duration for the single-form display yielding 80% recognition-memory accuracy. The obtained values were then used as a basis for matching *S*s in assigning them to either the one- or the three-syllable response group in the experiment proper. That this matching technique was effective in ensuring an equation in terms of visual perceptual ability is demonstrated by an obtained Pearson product-moment correlation of .86 between the recognition-memory exposure duration and the exposure duration determined for each *S* on the experimental task.

---

[6] F. Attneave, Physical determinants of the judged complexity of shape, *J. exp. Psychol.*, 53, 1957, 221–227.

[7] C. W. Eriksen, Temporal luminance summation effects in backward and forward masking, *Percept. & Psychophys.*, 1, 1966, 87–92.

[8] C. W. Eriksen, M. D. Pollack, and W. E. Montague, Implicit speech: A mechanism in perceptual encoding? *J. exp. Psychol*, 1970, in press.

Following the practice task and the assigning of *S*s on a matched basis to the one- or three-syllable response group, *S*s spent a session learning the respective response names for the eight experimental nonsense forms. The learning was accomplished by flash cards on which the nonsense form was on one side of the card and its name on the other. Practice was self-paced. The assignment of names to the eight nonsense forms was on the basis that the words correspond between the one- and three-syllable response groups; that is, a form called *hem* by *S*s in the one-syllable response group was called *hemisphere* by the *S*s in the three-syllable response group.

After *S*s had learned the names of the eight nonsense syllables to a criterion of twice through a flash-card set without error, they began a series of three additional practice sessions. They were presented the six-form experimental displays at a brief duration, with the probe occurring concurrently for 30 msec. The *S* was required to report the indicated form, and *E* provided feedback after every trial. During the last practice session, an exposure duration of the six-form display yielding approximately 85% accuracy with the concurrent probe was established for each *S*. This exposure duration was used for both the three- and the six-form displays for that *S* in the remaining experimental sessions. For the one-syllable group the mean duration was 3.4 msec.; it was 3.8 msec. for the three-syllable group.

Each *S* served eight experimental sessions. During each session he was presented with both three- and six-form displays and with a probe that occurred either concurrently with the display or was delayed by .5, 1, or 3 sec. Trials were organized in blocks of 12, and during an experimental session *S* received one block of trials at each combination of probe delay and display size. For each display size, each position and each form was indicated equally often over the eight sessions. The order of trial blocks within an experimental session was randomized. Each experimental session was preceded by a session with the flash cards to ensure that *S* was still capable of errorless performance in naming the nonsense forms. There were 5 min. of dark adaptation and warm-up practice (using the six-form displays with concurrent indicator) preceding the experimental trials. During the experimental sessions, feedback on accuracy was provided by telling *S* at the end of each trial block how many times he had been correct within that block.

RESULTS

In Figure 1 the average percentages of identifications that were

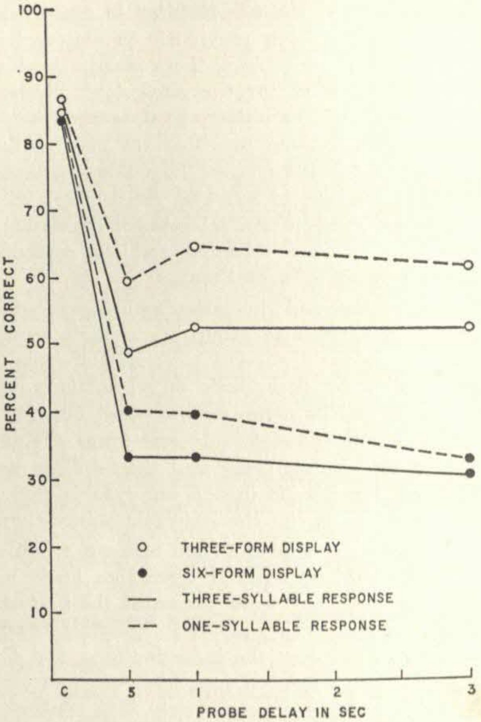


Fig. 1. Percentage of form identifications correct for the one- and three-syllable response groups as a function of number of forms in the display and delay of the probe

correct for the designated forms are shown as a function of the delay of the probe for the one- and three-syllable response groups. These data were analyzed in a matched-subject analysis of variance. Since the exposure duration of the display was varied for each subject so as to obtain approximately 85% correct performance when the probe occurred concurrently with the display, the concurrent probe data were left out of the analysis of variance. Thus only the data for probe delays of .5 sec. or longer were analyzed. As is apparent from Figure 1, performance for both three- and six-form displays was superior in the one-syllable response group [F (1, 4) = 78.09, $p < .01$]. For both groups performance was superior on the three-form displays, but this is a trivial and expected finding since the chance level of success is higher that the probe, when it occurs, designates one of the encoded forms for a three-form display as opposed to a six-form display.

Using the procedure of Estes and Taylor,[9] the three- and six-form displays were corrected for differences in chance level of performance and an estimate obtained as to the number of forms encoded by the subjects in the one- and three-syllable response groups. The formula is

$$PC = \frac{P}{D} + \frac{1 - P}{D} \times \frac{1}{8},$$

where PC is the obtained percent correct responses, P represents the number of forms effectively processed, D represents the number of forms in the display, and there are eight equally likely alternative responses. This procedure yielded estimates of 1.69 and 1.73 forms processed by the one-syllable response group on the three- and six-form displays respectively; for the three-syllable response group, the estimates of the number of forms processed were 1.32 and 1.36 for the corresponding display sizes. As is seen, there was little or no difference in the number of forms processed as a function of display size for either response group.

The analysis of variance revealed no significant effect for probe delay over the interval of .5 to 3 sec. [$F\ (2, 8) = 1.14, p > .20$], but the interaction of probe delay and display size obtained significance at the .05 level [$F\ (2, 8) = 6.05$]. The effect of subjects was also significant at the .05 level.

DISCUSSION

The superior performance of the one-syllable response group is consistent with the assumption that the noting or encoding process consists of implicit naming, or speech, responses. The longer the implicit naming response requires, the more the icon has decayed before the next item from the icon can be encoded, or named. However, there are considerations which suggest that the process and the explanation are not quite this simple. The superiority of the one-syllable group over the three-syllable group was not as great as would be anticipated on the basis of the time required to implicitly speak a one- as opposed to a three-syllable word.

We recorded the vocal durations of the one- and three-syllable words employed in the present experiment by feeding the output of a microphone into an oscilloscope. Spoken rapidly, the one-syllable words had a duration per word ranging from 200 to 250 msec., whereas the three-syllable words had a duration of approximately 500 msec.[10] Since implicit speech requires roughly the same durations as overt speech (see Landauer, n. 4), subjects in the one-syllable response group had, on the average, 200 msec. or more to encode a second form from the icon, or trace, of the display while subjects in the three-syllable response group were still in the process of encoding their first form. Given that the trace of the display persists for 500 msec., subjects in the three-syllable response group would have time to encode only one form whereas subjects in the one-syllable response group would be expected to encode two. But the obtained differences between the two groups were only about one-third this expected magnitude.

---

[9] W. K. Estes and H. A. Taylor, Visual detection in relation to the display size and redundancy of critical elements, *Percept. & Psychophys.*, 1, 1966, 9-16.

The discrepancy could be attributed to measurement error except that the data of Eriksen, Pollack, and Montague (see n. 8) indicate the same type of discrepancy. These investigators measured the latency from the presentation of a word or a two-digit number to the beginning of overt vocalization of the stimulus. They were testing the possibility that subjects encode a stimulus by implicitly speaking its name before overtly vocalizing. The assumption was supported to the extent that vocalization latencies were found to increase by 11 msec. per syllable for both words and number stimuli in their well-practiced subjects. But again, the latency differences did not correspond in magnitude to the differences in voicing duration for the different classes of stimuli. The latency for three-syllable words (the same as employed in the present experiment) was, on the average, 21 msec. longer than for one-syllable words (also the same as used in the present experiment). This despite the fact that it takes approximately 200 msec. longer to voice the three-syllable words.

Both the present results and the data from the study by Eriksen, Pollack, and Montague strongly suggest a relation between implicit speech and encoding of information in visual perception, even though the temporal transformation is not on a one-to-one basis. It is possible that encoding occurs in several stages. There may be a 'noting response,' some kind of attentional act, that precedes the implicit naming response which then encodes the information into a short-term storage system. This noting process may be able to operate several items ahead of the verbal encoding and, by having operated upon an item, thus delay its decay from the icon, or stimulus trace. The nature of this noting response may itself be an abbreviated form of an implicit naming response, which would account for the obtained relations between word length and encoding.

---

[10] The voicing duration of three-syllable words is not three times as long as that of one-syllable words because the two unaccented syllables do not take as much time as the accented syllable.

It would seem clear that in the present experiment the effect of response length operated in the encoding or noting process rather than in short-term memory. If the obtained results were to be attributed to different numbers of rehearsals in short-term memory, the performance difference between the one- and three-syllable response groups should have increased as the probe was delayed. This because the longer the interval between the display and the probe, the greater the disparity in the number of rehearsals that would have occurred. For the 3-sec. delay, one-syllable responses could have been implicitly spoken several times more often than the three-syllable responses. As the data showed, however, the performance difference was present when the probe occurred .5 sec. following the display, an interval too short to have permitted rehearsal in short-term storage.

The interaction between probe delay and display size is not readily explicable. For the six-form display, probe delay had little effect upon performance; but with three forms, performance was poorest when the probe occurred .5 sec. following the display. The most ready explanation is that the probe occurred at a point that interfered with the encoding process, but it is not apparent why this interference should have operated only for the three-form displays. However, this interaction was barely significant, and further work may reveal that it was attributable to chance factors.

STIMULUS SELECTION BY FIVE-, SEVEN-, AND NINE-YEAR-OLD CHILDREN

ROBERT L. SOLSO, *Loyola University, Chicago*

Abstract. Functional stimulus selection by five-, seven-, and nine-year-old children was examined in a paired-associate learning and recall test. The compound stimuli were of two classes—geometric figures and colors—and the responses were pictures of common objects. Five-, seven-, and nine-year-olds all seemed to select the geometric figures as functional stimuli. Several alternative hypotheses are offered to account for the results, among them an hypothesis of stimulus-habit hierarchy and an hypothesis of alternating stimulus selection.

The tendency to select a portion of a compound stimulus and to use that portion as a functional stimulus has been the object of considerable attention. Underwood and Schulz first suggested that stimulus selection plays an important role in paired-associate acquisition, especially when the stimulus elements elicit few associations.[1] In a direct test of this hypothesis, Underwood, Ham, and Ekstrand presented convincing evidence for stimulus selection as a function of the relative meaningfulness of the elements within compound stimuli.[2]

Recently Goulet, in an analysis of verbal learning in children, suggested that young children may tend to fractionate verbal stimulus terms more than older children or adults.[3] His rationale was that stimulus integration is a direct function of exposure, which generally corresponds to age. In addition, he offered an alternative suggestion, that differential acquisition rates, varying with age, may be due to non–stimulus selection among young children because of their naïveté in stimulus selection.

Received for publication October 22, 1969. An earlier part of the investigation was supported by PHS Research Grant MH-15722-01, National Institute of Mental Health. The author expresses his gratitude to B. J. Underwood, who generously commented on an earlier draft of the paper. Barbara S. Solso and Sheryl Stice assisted in the gathering of the data, and their contribution is gratefully acknowledged.

[1] B. J. Underwood and R. W. Schulz, *Meaningfulness and Verbal Learning*, 1960.

[2] B. J. Underwood, M. Ham, and B. Ekstrand, Cue selection in paired-associate learning, *J. exp. Psychol.*, 64, 1962, 405–409. The stimuli used in this study were colors and high- and low-meaningfulness trigrams.

[3] L. R. Goulet, Verbal learning in children: Implications for developmental research, *Psychol. Bull.*, 69, 1968, 359–376.

Although some investigators have used compound stimuli in learning studies across age-levels, a direct test of functional stimulus selection among children has not been actively examined. In the present study, children of five, seven, and nine years of age were presented with stimuli of two unique classes—colors and geometric forms—and with responses that were pictures of common objects. Following acquisition under compound-stimulus conditions, the children were asked to recall the response originally paired with one component of the compound stimulus. If during original learning, they selected form as a functional stimulus, then recall could be expected to be relatively faster for forms than for colors. On the other hand, if the children attended to both form and color during original acquisition, little difference between the recall rates of forms and colors would be expected. These factors were under direct investigation in the following study.

METHOD

General design. The *S*s learned a four-item paired-associate task in which compound stimuli were paired with a simple pictorial response. Following one perfect trial and a brief rest period, recall was tested in a series of transfer trials in which one-third of the *S*s received the original material, one-third were given one component of the compound stimulus paired with the responses from the first list, and the remaining third were given the other component of the compound stimulus paired with the original responses.

Subjects. A total of 180 *S*s, equally distributed among children five, seven, and nine years old at their last birthday, with an approximately equal number of boys and girls in each age-group, was used. Each original-learning cell consisted of 60 *S*s for each of the three ages studied. Since each age-group was subdivided during the transfer trials into recall conditions, each recall condition within age-groups consisted of 20 *S*s. The *S*s were selected from the public schools, Catholic schools, and college-campus schools, and were alternately assigned to test conditions.

Task and paradigms. The experimental task consisted of learning four paired associates using standard anticipation learning and then relearning four paired associates on a transfer task. Three different orders were used in original learning and transfer. Let A^1_2-B represent the learning paradigm in which compound stimuli (A^1_2) were employed. Following one perfect acquisition trial, 20 *S*s in each age group received one component in a transfer task (A_1-B), 20 *S*s received the other component (A_2-B), and the remaining 20 *S*s continued practice with the compound stimuli (A^1_2-B). If the functional stimulus is one component of the compound stimuli, then high positive transfer should be observed when that component is employed as a stimulus on the first trial for recall. For example, if five-year-old *S*s select A_2 as a functional stimulus from the A^1_2 compound, then the original learning paradigm is represented by A$_2$-B. The *S*s in the A$_2$-B recall condition should then manifest high positive transfer of learning relative to the *S*s in the A$_1$-B condition. Conversely, if A$_1$ serves as functional stimulus, then the transfer in A$_1$-B recall condition should be more rapid than in the A$_2$-B condition.

Transfer practice was continued until *S*s attained one perfect recall trial, but because of the possibility of differential learning rates with form stimuli as contrasted with color stimuli, only the first exposure to the stimulus components A$_1$, A$_2$, or A$_2^1$ was used in the analysis of stimulus selection.

Stimuli and procedure. The stimuli and responses were mounted on heavy, flat, gray cards measuring 5½ × 8½ in. The compound stimuli consisted of symbols surrounded by colors—red, yellow, blue, and orange. The color material was identical to the material used by Solso.[4] The color stimuli were 2½ × 3 in. cards affixed to the left portion of the gray cards. Within the color elements, nonglossy photographs of geometric symbols were placed. These symbols were reproduced on white paper 1½ × 2 in., thus leaving a ½-in. color frame around them when placed on the color elements. The symbols were a square, a diamond, a star, and a cross. The four response pictures consisted of approximations of items used in the early vocabulary of the Stanford Binet Intelligence Scale; namely, *horse, tree, telephone,* and *hat*. The *S*s responded with the name of the object pictured.

The instructions included a sample paired-associate card which contained a compound stimulus and response similar to the test material. The intent of the instructions was to identify the task without biasing the *S*s in favor of either stimulus. The test cards were manually presented using standard anticipation procedure, with a 3-sec. intratrial interval and a 6-sec. intertrial interval. One errorless trial was the criterion for learning. During recall conditions A$_1$-B and A$_2$-B, one stimulus was eliminated from the compound stimulus, as described above. The recall procedure included a 3:3 presentation rate for transfer trials.

RESULTS

The original-learning means (trials to criterion) indicated more rapid acquisition for the seven- and nine-year-olds ($\bar{X} = 6.9$, *SD* 3.7; $\bar{X} = 6.1$, *SD* 2.8) than for the five-year-olds ($\bar{X} = 10.6$, *SD* 6.5). An overall analysis of data yielded $F\ (2, 177) = 11.12$, $p < .001$.

The means and standard deviations for the recall conditions and the subjects' ages are presented in Table I and plotted in Figure 1. These data were treated by means of a 3 × 3 analysis in which one level consisted of age (five-, seven-, or nine-year-olds) and the other level consisted of recall condition (A$_1$, A$_2$, or A$_2^1$. The age-level yielded an $F\ (2, 171) = 5.07$, $p < .01$, while the recall condition yielded an

---

[4] R. L. Solso, Functional stimulus selection as related to color versus verbal stimuli, *J. exp. Psychol.*, 78, 1968, 382–387.

TABLE I

Means and Standard Deviations for First Transfer Trial

	Stimulus characteristics		
	Form alone	Color alone	Form and color
Five-year-olds			
$\bar{X}$	2.35	1.80	3.40
SD	1.32	1.12	.71
Seven-year-olds			
$\bar{X}$	2.75	2.20	3.45
SD	.38	1.00	.59
Nine-year-olds			
$\bar{X}$	3.20	2.45	3.60
SD	.78	1.02	.50

$F(2, 171) = 31.4$, $p < .001$. The F ratio for the interaction was not statistically significant.

Recall with color alone was the least effective condition, recall with form alone was more effective, and recall with the compound stimulus was the most effective. These results were consistent for all age-groups. It is apparent that appreciable recall occurred under all transfer conditions. Even the poorest recall—that of the five-year-olds, with color alone—was more than half as high as the same group's best recall, with the compound stimuli. This is to say that neither color nor form was selected exclusively by most of the Ss within a group.

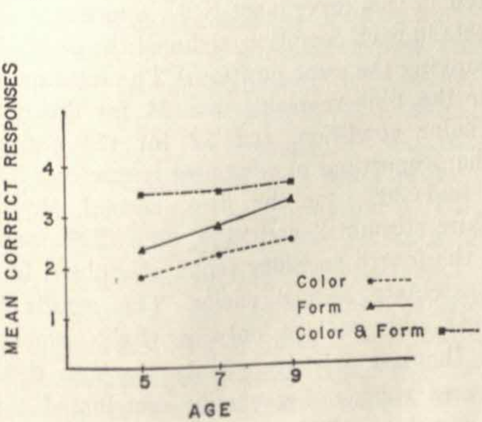


Fig. 1. Correct responses on first transfer trial for five-, seven-, and nine-year-olds

DISCUSSION

The present results indicated that functional stimulus selection does occur among five-, seven-, and nine-year-olds. The stimuli most frequently chosen by these groups were the geometric figures. But since there was a consistent and significant loss in recall of the geometric-figure stimulus as compared with the compound stimulus, it may appear that the figure component alone was not the *complete* functional stimulus. (An even greater disparity was noted between color stimuli and compound stimuli.)

This investigation also revealed significant factors due to age. Of particular interest was the convergence of the color-stimulus and form-stimulus conditions with the compound-stimulus condition during transfer as a function of age, even though the older subjects had less experience (trials to original learning) with the stimulus material than did the younger subjects. It is possible that the older subjects were better able to infer the last responses on the transfer trial because they had eliminated the other response alternatives. With only four pairs, subjects have an opportunity to deduce the last response with complete accuracy. What appears to have been increased stimulus selection for older subjects thus may, in reality, be a reflection of their ability to utilize a situation of ever-increasing constraint.

This possibility was elucidated by examining the relative number of correct responses on the first, second, third, and fourth transfer items for each age-group. (Obviously, the three different recall sequences used in this experiment, with approximately equal numbers of subjects in each condition, reduced the possibility of an easy response occupying the same position.) The total number of correct responses for the nine-year-olds was 64 for the figure condition, 49 for the color condition, and 72 for the compound-stimulus condition. The proportions of responses correct per item were 22%, 22%, 26%, and 30% for the first, second, third, and fourth positions (figure stimulus), and 21%, 26%, 22%, and 31% for the first through the fourth positions (color stimulus). In each case the last paired associate seemed easier. The results for the five- and seven-year olds yielded only a slight increase in correct responses for the last paired associate; in both these groups the correct responses were rather evenly distributed throughout the four paired associates. This analysis does indeed suggest that what seemed to be an increase in stimulus selection among the nine-year-olds may in large part have been due to their ability to use the decreasing sample of possible responses as a form of constraint and thereby deduce a correct response.

Several theoretical questions are raised by the results of this study. First, the results may be interpreted in terms of differential habit strength between the two stimuli and one response. Such an argument is consistent with a Hullian concept of response hierarchy, but in the current context it refers to a stimulus hierarchy and holds that the habit strength formed between A_2-B is greater than that formed between A_1-B. This concept has some similarity to a configuration hypothesis.

Second, the results may be interpreteted by suggesting that some subjects selected one stimulus of the compound while other subjects counterbalanced them by selecting the other component. (The mean recall, under these conditions, would not reflect the extensive selection which in fact occurred.) If this did occur, it should be reflected in an exaggerated standard deviation on the recall trial as a consequence of bimodality. Among the five-year-olds, form-stimulus recall condition, the standard deviation does appear to be exaggerated. In an attempt to further unravel this problem, it was reasoned that if one subject counterbalanced another subject by selecting different components, then this tendency would be manifest in a dichotomous frequency distribution of the raw recall data. These curves did not reveal convincing evidence for this hypothesis.

A third possibility carries the same rationale as the second possibility except that it suggests that the subjects may have selected one component of a compound stimulus on one paired associate and then selected the second component on another trial. If this had been occurring, then the possibility of a habit hierarchy as previously mentioned could have resulted. Furthermore, the results could be additionally complicated by a combination of the second and third possibilities. A final interpretation of the data is that subjects mediated from one stimulus element to another in recall. No evidence was found for or against this hypothesis.

The present experiment was not designed to test the specific mechanisms in operation during learning with a compound stimulus but rather to demonstrate the general tendency for young subjects to selectively attend to a cue within a compound. Further work will test specific hypotheses on stimulus selection.

RATE AND TERMINAL PROCESSES IN SKILL ACQUISITION

MARSHALL B. JONES, *Pennsylvania State University College of Medicine*

Abstract. After 16 trials of practice on micrometer adjustment, all subjects master the task, more or less; what little variance remains is mostly error. However, the rate at which subjects approach this terminal level of near-perfection can be retarded by systematically transforming knowledge of results: the stronger the transformation, the longer it takes subjects to master the task. With no such transformation, correlations among early trials of practice showed strong superdiagonal patterning. As practice proceeded, this pattern weakened and then, around Trial 7, collapsed; the correlation level dropped near zero and the matrix became disorganized. With progressively stronger transformations of knowledge of results, and in exact proportion as learning was retarded, the superdiagonal pattern persisted longer and longer into the practice series before it finally disintegrated. These results are interpreted in terms of a two-process theory of individual differences in motor learning, based on the distinction between the level of proficiency a subject approaches at the end of practice and the rate or route he follows in reaching it.

Extended practice on any skill necessarily results in either of two general circumstances. Either all subjects approach a common level of proficiency, or they approach different levels of proficiency. In the former case, performance may be adulterated by varying amounts of random responding, and in the latter case, the terminal levels may be more or less different. Still, a good deal depends on whether practice brings the subjects to a common level of skill or spreads them out into stable but different levels. Practice on complex tasks generally spreads the subjects out, while practice on a few very simple tasks—lever positioning, line drawing, micrometer adjustment—brings them together, because anybody can learn to master these tasks. Some take longer than others, but eventually everyone performs to near-perfection.

In most cases, the correlations between trials of practice follow the *superdiagonal pattern.*[1] The correlations are largest between neighboring trials and become smaller as the trials are more and more separated—become smaller, that is, across the rows to the

---

Received for publication October 20, 1969. The research reported was sponsored by the Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force, under Grant F44620-68-C-0072.

[1] M. B. Jones, Molar correlational analysis, *USN Sch. Aviat. Med. Monogr.*, 1960.

right or up the columns. The smallest correlation is in the upper right-hand corner of the matrix. This pattern is the rule in matrices of intertrial correlations,[2] but there are exceptions. In a few very simple tasks—in lever positioning, line drawing, micrometer adjustment; those same tasks in which all subjects approach a common level of near-perfection—only vestiges of the superdiagonal pattern may be seen.[3] In one task, lever positioning, the pattern appears to be present at first and then to degenerate as learning progresses.[4] The evidence, moreover, suggests that the association with learning is essential. As all subjects approach perfection and as reliable differences in performance are eroded away, error elements loom larger and larger until the correlation level finally drops toward zero and the pattern becomes completely disorganized.

In lever positioning, the erosion of reliable differences and the concurrent degeneration of the superdiagonal pattern appears clearly only in blocks of interpolated trials, when the subjects are given no knowledge of results. In micrometer adjustment, however, the degeneration of the superdiagonal pattern does take place with knowledge of results given on every trial, and it does so in a highly systematic way: the superdiagonal pattern is present at first and then decays with practice. Micrometer adjustment, therefore, affords us an opportunity not only to study the learning process in the absence of terminal differences in performance but to study it with particular ease and directness.

METHOD

Subjects. The *Ss* were 200 airmen in basic training at Lackland Air Force Base. None had previous experience with the micrometer, and all had completed at least eight years of formal schooling.

Apparatus. The apparatus was a micrometer mounted so that its turning knob projected through a wooden screen; the micrometer scale was visible only to *E*. A single full turn of the knob counted 25 turning units. The task the *Ss* were to learn was to turn the knob 200 units or eight full turns. They knew the goal was 200 units, but not that one turn equalled 25 units. (In turning the knob a standard grip and turning motion were required: thumb and tip of forefinger on the knob, and wrist motion.)

---

[2] M. B. Jones, Practice as a process of simplification, *Psychol. Rev.*, 69, 1962, 274-294; M. B. Jones, Individual differences, in E. A. Bilodeau (ed.), *Acquisition of Skill*, 1966.

[3] M. B. Jones, Differential processes in acquisition, in E. A. Bilodeau (ed.), *Principles of Skill Acquisition*, 1969.

[4] M. B. Jones, Knowledge of results and intertrial correlations in a simple motor task, *J. mot. Behav.*, 1, 1969, 331-340.

Procedure. The *S*s were divided into five groups of 40. The groups differed in the feedback they received from *E* after each trial. In Group 100 *S*'s response was reported truly. If he turned the knob four full turns, or 100 units, his reported score was 100 units. Since *S* knew that the goal was 200 units, he also knew that he was 100 units short. In the other four groups, *S*'s true response (X) was systematically transformed into his reported score (Y). The transformations were: for Group 75, $Y = 50 + 0.75X$; for Group 50, $Y = 100 + 0.50X$; for Group 25, $Y = 150 + 0.25X$; and for Group 12.5, $Y = 175 + 0.125X$. In Group 50, for example, an *S* who turned the knob 100 units was told that he had turned it 150 units.

All *S*s were given 16 trials, each of which was followed by a reported score. About 5 sec. elapsed between the response and the reported score; 10 sec. elapsed between the giving of the score and the signal to turn the knob again. In each trial, the dependent measure was *S*'s absolute deviation in true (untransformed) units from the target.

RESULTS

This experiment was conceived and carried out by E. A. Bilodeau.[5] The trial means in the five groups were reported by him in his original report[6] and are reproduced in Figure 1. All five groups started at approximately the same point, just under 180 deviation units from the target of eight full turns; and they all ended with mean absolute deviations around 13 units. The rates, however, at which the response was acquired varied systematically. The curve for Group 100 was negatively accelerated throughout and reached asymptotic levels around Trial 9, long before practice ended; learning was rapid. In Group 75 it took the subjects somewhat longer to reach asymptotic values. In Group 50 it took them longer yet, and the performance curve was definitely sigmoid. These trends continued in Groups 25 and 12.5. In Group 12.5, learning was slowed to the point that the mean just reached asymptotic levels by Trial 16.

Table I presents the intertrial correlations plus means and standard deviations for Group 100. The table is broken between

---

[5] The author is indebted to E. A. Bilodeau and to Ina McD. Bilodeau for their permission to report these data here. In the original report of this work, the data were analyzed with respect to group differences only, not individual differences.

[6] E. A. Bilodeau, Speed of acquiring a simple motor response as a function of the systematic transformation of knowledge of results, this JOURNAL, 66, 1953, 409–420.

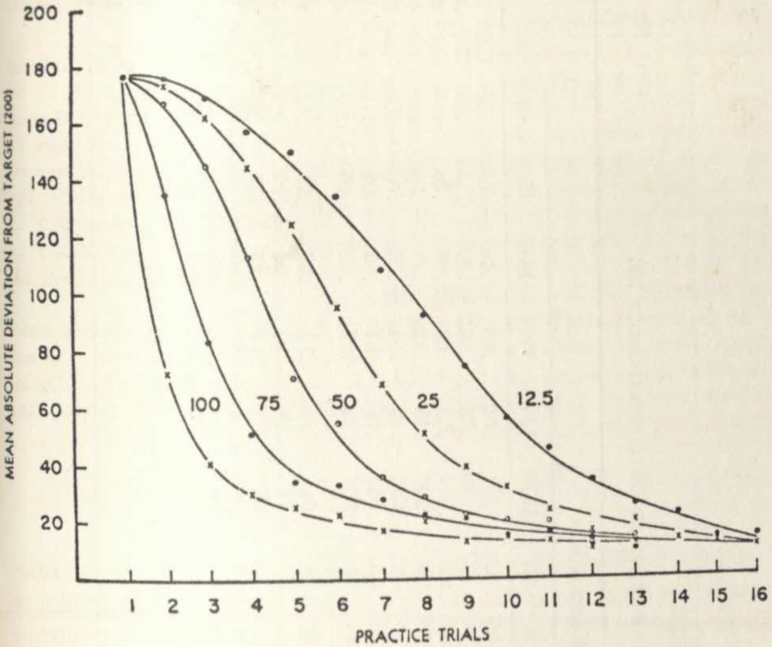


FIG. 1. Mean absolute deviation of true scores from goal score, plotted against successive trials

Trials 7 and 8 in order to bring out the most striking feature of the results. In Trials 1–7 the intertrial correlations shape themselves into a superdiagonal form of moderate level and substantial regularity. From Trial 8 on, however, the pattern disappears; the correlations among Trials 8–16 show hardly a trace of superdiagonal form; and the cross-correlations in the rectangular block are almost completely disorganized.

The correlational results for Group 75, which appear in Table II, are much the same, except that the superdiagonal form remains intact through Trial 9; the correlations among Trials 10–16 show a weak tendency toward superdiagonality.

The correlational results for Groups 50, 25, and 12.5 are presented in Tables III, IV, and V respectively. All three tables show the same general features as do the correlations for Groups 100 and 75, except that the superdiagonal form persists longer in each successive group. In Group 50 the break does not come until Trial 12; in Group 25 the pattern is still strong after Trial 13, is

TABLE I

INTERTRIAL CORRELATIONS, MEANS, AND STANDARD DEVIATIONS: GROUP 100

Trials	Intertrial correlations																$\bar{X}$	σ
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16		
1		.30	.33	.29	.12	.14	−.16	.07	.05	.10	−.00	.04	.03	−.16	−.06	.05	178	7
2			.72	.56	.43	.20	−.10	.34	.11	.15	.22	.08	.01	−.02	.13	.06	73	54
3				.72	.66	.47	.15	.34	.17	.31	.20	.03	−.08	.06	.03	.18	41	42
4					.57	.57	.26	.08	.06	.29	.20	.10	.13	.13	−.05	.20	32	44
5						.77	.44	.14	.23	.48	.09	−.10	−.04	−.05	.00	−.04	26	32
6							.65	.15	.30	.62	−.01	.11	.01	.02	−.09	.14	23	30
7								.19	.36	.60	−.08	.22	.17	.13	.09	.23	18	17
8									.19	.09	−.04	.28	.33	.27	.27	.42	20	18
9										.47	.14	−.08	.20	.04	.12	.05	13	11
10											.08	−.00	.19	.20	.27	.20	15	15
11												−.16	.17	.09	.19	−.03	14	9
12													.09	.38	.20	.41	11	9
13														.09	.06	.12	13	12
14															.52	.31	14	11
15																.22	12	9
16																	11	9

TABLE II

Intertrial Correlations, Means, and Standard Deviations: Group 75

	Intertrial correlations																	
Trials	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	$\bar{X}$	σ
1		.58	.42	.42	.40	.19	.12	.08	.13	−.01	.02	.12	.11	.44	.15	.35	178	6
2			.64	.66	.46	.25	.04	.13	−.01	−.01	−.14	.01	−.19	.16	−.12	−.02	135	26
3				.78	.58	.41	.24	.25	.17	.11	−.06	−.12	−.10	.23	−.20	−.01	84	44
4					.68	.49	.42	.43	.30	.03	.04	.07	−.00	.15	−.15	−.18	52	49
5						.78	.63	.64	.57	.32	.12	.23	.22	.21	−.01	−.02	35	43
6							.73	.66	.64	.35	.02	.30	.26	.16	−.06	−.03	34	35
7								.80	.61	.15	.21	.27	.27	.20	.02	−.00	28	39
8									.68	.32	.30	.37	.22	.07	−.19	.08	22	25
9										.36	.31	.20	.27	.21	−.18	.04	22	19
10											.24	.43	−.03	.13	−.06	.14	15	13
11												.07	−.16	.07	−.12	−.23	16	16
12													.22	−.03	−.16	.20	14	12
13														.13	.20	.23	11	10
14															.06	.10	14	10
15																.21	16	15
16																	13	11

TABLE III

INTERTRIAL CORRELATIONS, MEANS, AND STANDARD DEVIATIONS: GROUP 50

Trials	Intertrial correlations																$\bar{X}$	σ
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16		
1		−.75	.61	.47	.33	.34	.30	.32	.22	.16	.03	.18	.16	−.08	.28	.08	179	7
2			.85	.68	.51	.52	.44	.44	.35	.35	.13	.28	.10	.09	.19	.06	167	14
3				.71	.57	.53	.44	.38	.30	.30	.13	.30	.07	.06	.17	.05	145	29
4					.79	.65	.54	.47	.37	.33	.11	.38	−.13	.01	.06	−.06	116	44
5						.71	.70	.61	.54	.42	.20	.36	.02	.23	.17	−.08	72	53
6							.85	.72	.69	.51	.36	.47	.03	.25	.22	.05	55	56
7								.90	.79	.61	.38	.49	.01	.34	.21	−.10	36	45
8									.86	.68	.49	.49	.12	.44	.16	−.05	29	39
9										.78	.71	.66	.16	.54	.28	.08	23	33
10											.81	.73	.22	.66	.27	.14	21	28
11												.74	.30	.71	.34	.22	20	26
12													.16	.56	.29	.02	13	16
13														.44	.42	.36	15	12
14															.39	.24	13	11
15																.60	14	14
16																	13	10

TABLE IV

INTERTRIAL CORRELATIONS, MEANS, AND STANDARD DEVIATIONS: GROUP 25

Trials	Intertrial correlations																$\bar{X}$	σ
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16		
1		.81	.64	.49	.46	.36	.33	.36	.37	.37	.20	.22	.19	.17	−.09	−.06	177	7
2			.79	.59	.48	.44	.39	.34	.31	.33	.20	.25	.25	.26	−.06	−.16	173	8
3				.89	.77	.74	.59	.47	.32	.40	.13	.20	.24	.37	−.19	−.25	162	16
4					.88	.83	.69	.55	.35	.38	.13	.21	.24	.44	−.06	−.10	145	31
5						.85	.75	.63	.43	.41	.22	.24	.27	.47	.04	−.04	125	46
6							.90	.66	.47	.50	.27	.26	.27	.51	.03	−.06	95	57
7								.80	.63	.60	.42	.37	.27	.51	.07	.04	68	55
8									.76	.75	.52	.46	.27	.45	.08	.05	51	48
9										.78	.58	.50	.42	.28	.13	−.01	39	41
10											.53	.54	.44	.31	−.02	−.25	33	42
11												.73	.54	.27	.18	.09	24	28
12													.78	.38	.18	.08	18	23
13														.48	.19	−.03	20	21
14															.21	.05	15	15
15																.41	13	13
16																	14	13

TABLE V

Intertrial Correlations, Means, and Standard Deviations: Group 12.5

	Intertrial correlations																	
Trials	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	X	σ
1		.69	.62	.56	.45	.40	.33	.36	.45	.37	.40	.41	.21	.26	.17	.18	178	7
2			.83	.75	.82	.67	.45	.37	.41	.37	.36	.31	.29	.23	.23	.24	176	12
3				.91	.86	.74	.50	.48	.43	.41	.33	.22	.26	.20	.16	.24	170	19
4					.84	.79	.58	.58	.53	.44	.39	.27	.24	.23	.12	.26	157	33
5						.90	.66	.61	.54	.51	.39	.32	.34	.25	.25	.21	149	34
6							.81	.80	.71	.64	.43	.44	.40	.33	.33	.18	134	40
7								.84	.74	.72	.56	.55	.45	.35	.28	.23	108	54
8									.83	.77	.53	.62	.58	.42	.32	.21	93	58
9										.83	.66	.66	.51	.46	.45	.09	75	58
10											.83	.79	.65	.52	.48	.16	56	52
11												.84	.66	.51	.44	.30	46	50
12													82	.69	.63	.34	34	39
13														.69	.65	.40	27	29
14															.66	.30	23	25
15																.29	15	12
16																	14	9

equivocal at Trial 14, and breaks down entirely in the last two trials. In Group 12.5 the pattern does not disintegrate until the last trial, and even there the correlations are positive, though they do not increase toward the superdiagonal.

Dividing the trial sequences into initial (superdiagonal) and terminal (post-superdiagonal) segments may be misleading. It does not mean either that the superdiagonal segment is without irregularities or that the post-superdiagonal segment is completely disorganized. In Table III, for example, the correlations among Trials 13–16, the post-superdiagonal segment, show reasonably good superdiagonal form. The differential process that accompanies practice is continuous; it does not break off in sharp discontinuities. Nevertheless, it does become plain at some point that the superdiagonal pattern no longer organizes the correlations. The main finding—and the one that dividing the trial sequences was intended to underscore—is that this point comes later and later in practice as we turn from Group 100 to Groups 75, 50, 25, and 12.5.

The standard deviations are also noteworthy. In all groups the standard deviation of Trial 1 was around 7 units and the standard deviation of Trial 16 around 10. In Group 100 the largest standard deviation was at Trial 2; after Trial 2 the standard deviations took a generally decreasing course, becoming asymptotic around 9 units at Trial 11. In Group 75 the standard deviations increased from Trial 1 to a maximum value of 49 at Trial 4 and then decreased by stages to asymptotic levels. In Group 50 the 'peak' in the standard deviations came two trials later, at Trial 6. In Group 25 the standard deviations also peaked at Trial 6, though the value for Trial 7 was almost as large. In Group 12.5 the peak came at Trials 8 and 9. In all groups the standard deviations started small, increased to a maximum, and then fell away to asymptotic levels; but the trial at which the maximum value occurred came later and later in the trial sequence, as the acquisition process was slowed down by progressively stronger transformations of knowledge of results.

DISCUSSION

In micrometer adjustment, differential organization decays in precise proportion as learning advances. Groups 100, 75, 50, 25, and 12.5 reached asymptotic mean performance later and later in practice. Meanwhile, the intertrial correlations started out strongly patterned in all five groups; then the pattern weakened and broke down in the same order as the means became asymptotic. The pattern broke down first in Group 100 and then, in order, in Groups 75, 50, 25, and 12.5. There was a perfect association between mean mastery of the task and disorganization of the correlation pattern. As reliable differences in performance shrank and then vanished, as all subjects approached the same level of near-perfection, the differential organization of the task weakened and then collapsed.

The behavior of the trial variances is also noteworthy. On Trial 1 the subjects were asked to turn the knob around once (25 units); thereafter, they were given knowledge of results. Since the goal was 200 units, or eight full turns of the knob, the subjects tended to undershoot in the next few trials. Early in practice the mean response (in raw, not deviation, units) increased as the subjects worked their way from one to eight full turns of the knob. As the mean response increased, the variance associated with it also increased. Then, as the subjects began to 'home in' on the goal, the variance started to drop; by the end of practice the standard deviations were uniformly less than 15 units, like the means.

In all groups, then, there was an initial rise in variance followed by progressive shrinkage, although the groups differed in the rates at which they ran this course. The differences, however, fell into the same order as the mean and correlational differences. As mean performance slowed up from group to group, as the superdiagonal pattern broke down progressively later in practice, the peak variance came later and later in the trial sequence.

The importance of these results depends on the fact that in micrometer adjustment, terminal differences in performance are absent. The correlaton patterns that appear in Tables I–V reflect the rates at which different individuals approached a common level of proficiency. What we have in these tables is the *rate process* uncontaminated by terminal differences. When everyone terminates at the same level of proficiency, the superdiagonal pattern runs a weakening course with practice, correlation level drops, and regularity degenerates progressively, until finally the pattern becomes completely disorganized with correlations ranging positively and negatively around zero. This gradually disorganizing superdiagonal pattern is a specific expression of the rate process.

By the *terminal process* in practice, I mean the levels of proficiency different individuals achieve after unlimited practice. In micrometer adjustment there is no terminal process as process—only a common termination. Hence, the correlation patterns in Tables I–V are pure rate patterns. It would be useful if we had pure terminal patterns as well, but it is more or less certain that we never will, because people do not proceed toward different terminal levels by the same routes, with no individual variation as to rate or course. Fortunately, our inability to isolate the terminal pattern poses no great obstacle, because we know what this pattern must be. The correlations generated by the terminal process are controlled by a single common factor, level of proficiency after unlimited practice; they form a Spearman unit hierarchy and obey the double-tetrad rule.[7] The loadings, moreover, on this one common factor increase monotonically with practice; the first trial has the smallest loading, the second a larger one, the third still larger, and so on, while the last trial has the largest loading. Hence, the correlation pattern is a foregone conclusion. The correlations increase down the columns and across the rows to the right.

The factor which rules the terminal process is not hypothetical in the usual sense. It is true that we do not know *exactly* what the terminal levels of proficiency are; but we do have a series of empirical variables—the trials of practice—which converge on the terminal measure. In the case of micrometer adjustment, of course, there is no convergence because there is no terminal process. However, in more complex tasks, where practice spreads people out into different and stable terminal levels, trials of practice converge quickly and closely on the theoretical terminus. Intertrial correlations late in practice (in most tests, after 16 trials) are typically in excess of .90; and loadings on the one common factor, while they never equal unity, at least theoretically, come very close to it.

With the distinction between rate and terminal processes, we have arrived at a two-process theory of individual differences in learning. According to this theory, individual differences in learning should be understood in terms of (*a*) a terminal process that is controlled by final levels of proficiency and (*b*) a rate process that reflects individual differences in the routes (the courses) different individuals take in arriving at their terminal positions. The terminal process generates a Spearman unit hierarchy with loadings on the one common factor that increase monotonically with practice; the rate process generates a superdiagonal form that is strong at the beginning and wastes away with practice. The total correlation pattern is a combination of these two component patterns and should be analyzed in terms of them.

---

[7] C. Spearman, General intelligence, objectively determined and measured, this JOURNAL, 15, 1904, 201–293.

The principal alternative to this theory is factor analysis. Many psychologists understand individual differences in learning (as in every other area) in terms of hypothetical factors defined by factor-analytic convention. In the psychomotor area these factors depend partly on the trials of practice and partly on reference tests administered to the same subjects. Thus, the same trials of practice might have a different factorial composition if they were administered in conjunction with different reference tests.

The two-process theory differs from the factor-analytic approach in two fundamental respects. First, the two-process analysis is *empirical;* it involves no hypothetical factors, except to the extent that the last trial of practice fails to converge on final levels of proficiency. Second, the two-process analysis decomposes the differential process that accompanies acquisition into *empirically separable components*. Factor-analytic factors are purely analytic; they appear nowhere in pure form but only in conjunction with other factors. Theoretically they account for part of the covariance between trials, but the part that one factor is supposed to explain never appears uncontaminated by other parts attributed to other factors. In the two-process theory, on the other hand, the rate process is empirically separable from the terminal process. The whole importance of micrometer adjustment centers in the fact that in this simple task there are no differences in terminal proficiency. Hence, all reliable differences between subjects are due to differences in rate.

The ordinary psychomotor situation involves both rate and terminal processes; and in general, the total pattern is superdiagonal. However, the separability of the two components suggests that there may be an exception to this rule. Terminal and rate patterns do not differ in the columns; in both cases the correlations increase toward the superdiagonal. However, they do differ along the rows. In the rate process the correlations increase to the *left*, while in the terminal process they increase to the *right*. Now, in the early stages of practice the rate process is at its strongest while the terminal process is weak. Hence, in the early stages of practice the total correlations will very probably decrease to the right, in the superdiagonal way. However, late in practice the rate pattern has all but wasted away and the terminal pattern is at its strongest. Therefore, it is possible that late in practice the two components may combine to produce a total pattern which is not superdiagonal. If individual differences in rate peter out *before* most of the subjects have reached their terminal positions—if the terminal process continues a while after the rate process has stopped, or to put it another way, if individuals late in practice converge on their terminal positions in a uniform fashion—then we should see a piece of the Spearman unit hierarchy 'sticking out' past the superdiagonal pattern. Late in practice the correlations should increase to the *right* along the rows, instead of to the left.

In the two-process theory, rate and terminal processes are not analytic artifacts but independent sources of covariance in the intertrial matrix. In consequence, it is possible for trials late in practice to show features (along the rows) of the Spearman unit hierarchy rather than the superdiagonal pattern, in violation of the general rule that correlations between trials of practice, particularly in complex tasks, follow the superdiagonal pattern. Since nothing similar is predicted by the factor-analytic approach, a review of known intertrial matrices would seem to be in order. If any one of these matrices shows a clear tendency for correlations late in practice to increase along the rows to the right, the fact would constitute strong confirmatory evidence for the two-process theory.

SUMMARY

Five groups of subjects practiced micrometer adjustment; knowledge of results was subjected to progressively stronger transformations from one group to the next. The effect of these transformations, as reflected in the trial means, was progressively to retard learning over the 16-trial practice sequence. As the transformations became stronger, it took the subjects longer and longer to achieve mastery over this very simple task; in the end, however, all subjects in all groups reached near-perfect levels of performance. Meanwhile, the patterning of intertrial correlations underwent parallel changes. In all groups the early trials showed clear superdiagonal form, but this pattern persisted later and later in the trial sequence as the transformations became stronger and stronger, although ultimately correlation level dropped drastically in all groups and the correlations became disorganized.

The interpretation of these results hinges on the absence of reliable individual differences after 16 trials of practice. As practice proceeds, error variations play a larger and larger role. In the early trials there are reliable individual differences in the rates or courses different individuals take in reaching the same terminal level, and these differences generate strong superdiagonal form. However, as practice proceeds, and the subjects converge more and more on a common level of terminal proficiency, reliable differences among them shrink steadily until finally little, if anything, remains except random error. At this point the intertrial correlations drop toward zero and all pattern disintegrates. This entire series of changes—the early appearance of superdiagonal form, its gradual weakening and final collapse, together with the reasons therefor—is called the *rate process,* because it grows out of differences in the rates or routes that different individuals follow in acquisition rather than out of the levels of proficiency in which they terminate. In more complex and difficult motor tasks, there are substantial and abiding differences in terminal performance as well, and a *terminal process* that grows out of these differences. The terminal process generates a Spearman unit hierarchy in the intertrial correlations with loadings on the one common factor that increase monotonically through the practice sequence. The total correlations among trials of practice should be understood and analyzed as combinations of these two processes and the correlation patterns they generate.

THE RELATIONSHIP OF STATE ANXIETY, FEEDBACK, AND ONGOING SELF-REPORTED AFFECT TO PERFORMANCE IN COMPLEX VERBAL LEARNING

C. R. Snyder and Martin Katahn, *Vanderbilt University*

Abstract. The feedback in this study consisted of comparisons with high, average, and low standards of achievement; the verbal-learning task correlated .40 with actual classroom learning. Higher anxiety was found to be significantly related to poorer performance on the task. Going from high, to average, to low standards of comparison, there was a significant trend for performance to improve. For the group as a whole, positive self-reported affects correlated positively with performance and negative affects correlated negatively. For the medium- and low-anxiety groups, a majority of the positive affects correlated positively and a majority of the negative affects correlated negatively. For the high-anxiety group, however, a majority of the positive affects correlated negatively and a majority of the negative affects correlated positively with performance. Affect change preceded establishment of any significant affect-performance correlations. That the significant effects of feedback and anxiety were not reflected in the self-report scales of affect highlights the importance of verifying the outcome of experimental manipulations assumed on an a priori basis to alter the subject's internal state.

The purposes of this study were to investigate the relationship of state anxiety, feedback consisting of comparisons with high, average, and low standards of achievement, and ongoing affective experiences to performance on a complex learning task. The experimental variables and methodology were selected after consideration of several lines of research.

Since the development of the Manifest Anxiety Scale and the Test Anxiety Questionnaire, interest in the relationship of affect to verbal learning has centered predominantly on anxiety.[1] Results of numerous studies suggest that stronger and more consistent relationships are obtained when specific measures of situational anxiety (such as the Test Anxiety Questionnaire) are correlated with performance than when a general measure (such as the Manifest Anxiety Scale) is employed.[2] Consistent with these findings,

Received for publication November 11, 1969.

[1] Janet A. Taylor, A personality scale of manifest anxiety, *J. abnorm. soc. Psychol.*, 48, 1953, 285–290; G. Mandler and S. B. Sarason, A study of anxiety and learning, *J. abnorm. soc. Psychol.*, 47, 1952, 166–173.

[2] I. G. Sarason, Empirical findings and theoretical problems in the use of anxiety scales, *Psychol. Bull.*, 57, 1960, 403–415.

subjects for the current study were selected according to their scores on a 'state' form of the State-Trait Anxiety Inventory (STAI), designed to estimate feelings before a classroom test.[3] The STAI has been evaluated as the most sophisticated of available scales for measuring anxiety.[4] The first purpose of the study reported here was to test the most recent anxiety-theory formulations, using a situationally specific measure of anxiety.[5]

A second purpose of this study was to ascertain the effect on performance of feedback that consisted of either high, average, or low scores, each type being presented to different groups as the 'mean' attained by subjects in a previous experiment at the university. Subjects comparing themselves with a high standard—one generally attained by 17% of the students in an earlier sample—were expected to experience more negative affect. Subjects comparing themselves with a low standard—one exceeded by 83% of the earlier sample—were expected to experience less negative affect. Other research using 'failure feedback' has obtained an interaction effect of anxiety level and failure. Overall, high-anxiety subjects have been inferior to low-anxiety subjects. In almost all reported studies, the high-anxiety subjects have suffered a decrement in performance following failure feedback, in comparison with the high-anxiety nonfailure groups. For the low-anxiety subjects no decrement in performance has followed failure feedback, in comparison with nonfailure groups.[6]

Considering the emphasis which theorists have placed on affect as a motivational source, there has been relatively little research relating affects other than anxiety to performance.[7] One area of research receiving scant attention is the relation of affect to performance at the time of the task, to performance on that task. Katahn, Snyder, and Durlak investigated self-reported affect during performance in serial learning, noting the greater the degree of negative affect, the poorer the performance.[8] The third purpose of

[3] C. D. Spielberger and R. L. Gorsuch, Mediating process in conditioning, Report of United States Public Health Service Grants MH-7229, MH-7446, and HD-947, 1966.

[4] E. E. Levitt, *The Psychology of Anxiety,* 1967, 71.

[5] K. W. Spence and Janet A. Spence, Motivational components of manifest anxiety: Drive and drive stimuli, in C. D. Spielberger (ed.), *Anxiety and Behavior,* 1966, 291–326.

[6] Spence and Spence (see n. 5 above), 312.

[7] C. E. Izard and S. S. Tomkins, Affect and behavior: Anxiety as a negative affect, in Spielberger (see n. 5 above), 81–123.

[8] Martin Katahn, C. R. Snyder, and Joseph A. Durlak, The relationship of self-reported affect to performance in serial learning, *Psychon. Sci.,* in press.

the current study was to examine the relationship of an expanded self-report scale (including both positive and negative affects), to performance on a complex verbal-learning task.

The complex verbal-learning task used in this experiment was designed to have a reasonable correlation (about .40) with 'real-life' classroom learning.[9] The subject's job was to learn basic concepts from Hullian theory. Thus, by the nature of the task employed in this study, fairly strong inferences from the laboratory to naturally occurring situations might be made.

METHOD

Subjects. Ninety men, students at Vanderbilt University, served in the experiment as part of their introductory-psychology course requirements. *S*s were selected according to their scores on a 'state' form (feelings before a classroom test) of the STAI, with 30 chosen from the top 20%, 30 from the middle 40%, and 30 from the lowest 20%. The STAI was given at the beginning of the semester, which was one month before the start of the experiment. The anxiety variable consisted of low, medium, and high levels of anxiety (LA, MA, and HA).

Materials. The learning task, described more fully elsewhere, consisted of 16 items covering Hullian learning theory and 14 questions based upon the information items.[10] For example, one information item was: "Conditioning is the process by which a stimulus comes to elicit a response it did not formerly elicit." The related question was: "__________ is the process by which a stimulus comes to elicit a response it did not formerly elicit." Similarly, all questions left a blank for the appropriate answer. All information and question items took up two lines of space on slides, and were presented visually on a screen for groups of *S*s.

Design and procedure. The design was a 3 × 3 × 2 factorial, the variables being anxiety, feedback, and report versus no report of affective experiences. Groups of approximately 15 *S*s were administered the task in a classroom. Each *S* had a booklet with an answer sheet for each trial.

Feedback was manipulated at three levels: negative, average, and positive. The 'average number' of correct responses for Vanderbilt students appeared at the top of the answer sheet for each trial. The negative-feedback condition consisted of 'average numbers' of correct responses, numbers which were in reality one standard deviation above previously established means for each trial. The average-feedback condition consisted of numbers of correct responses which were the same as previously established means. The positive-feedback condition consisted of numbers of correct responses which were one standard deviation below previously established means. A majority of the *S*s in the negative-feedback condition were expected to fall below the standards, thereby experiencing more negative affect. A majority of the *S*s in the positive-feedback condition were expected to exceed the standards, thereby experiencing more positive affect. The *S*s in the average-feedback condition were expected to fall between the other two feedback conditions with respect to affect.

---

[9] Martin Katahn and Lee Branham, Effects of manifest anxiety on the acquisition and generalization of concepts from Hullian theory, this JOURNAL, 81, 1968, 575–580.

[10] Katahn and Branham (see n. 9 above).

The report variable consisted of two conditions, groups receiving (report condition) and groups not receiving (no-report condition) the scales of self-reported affect. The scales consisted of the words *annoyed, interested, anxious, calm, unpleasant, satisfied, bored, pleasant, excited, relaxed, tense, discouraged,* and *ashamed.* The *S* could vary his response from 1 to 9 (with 1 denoting 'slightly,' 5 'moderately,' and 9 'strongly') to show the degree to which he felt each word was descriptive of his experience at the moment. A self-report scale was taken before each trial for the *S*s in the report condition. Two sets of instructions were necessary. Subjects in the report condition were told they would fill out a short questionnaire describing how they felt at that moment in the experiment, would study information presented via slides on the screen, and then would supply fill-in-the-blank answers to questions over the information. The instructions were the same for the no-report condition, except the instructions about the self-report scales were omitted.

Six trials were given. For *S*s in the report condition, each trial consisted of (*a*) allowing the *S* 60 sec. to fill out the self-report scales, (*b*) presenting the 16 information items, each visable for 8 sec., and (*c*) presenting the 14 questions, each visable for 8 sec., with instructions to write the answer. The no-report condition differed only in that the *S*s rested for 60 sec. before each trial instead of filling out the self-report scales. The order of information items was constant over all six trials, while the questions were randomized so as to minimize serial learning.

RESULTS

Considering anxiety, feedback, and report versus no-report as between-subject variables, and trials as a within-subject variable, scores for the six trials were analyzed in a four-way analysis of variance. The only significant between-subjects variable was a main effect of anxiety ($F = 4.79$, $df = 2/72$, $p < .01$). As Figure 1 shows, the low-anxiety subjects in the report condition (LA-R) and in the no-report condition (LA-NR) were superior to the medium-anxiety subjects in both report and no-report conditions (MA-R and MA-NR), who were in turn superior to the high-anxiety subjects in both report and no-report conditions (HA-R and HA-NR). There was a trend for the more positive feedback to be associated with better performance, although this trend did not reach statistical significance ($F = 2.02$, $df = 2/72$, $p < .13$). A nonparametric sign test comparing each of the three feedback conditions at each anxiety

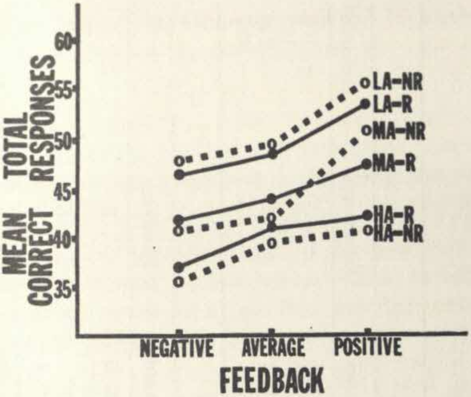


FIG. 1. Mean number of total correct responses as a function of anxiety, feedback, and report versus no-report conditions

level and report condition was performed. For each anxiety level and report condition, the positive-feedback condition was superior to the average-feedback condition, and the average-feedback condition was superior to the negative-feedback condition. Thus, in comparisons of 12 possible pairs of conditions, the more positive feedback was associated with better performance in every case, and this trend was significant at the .001 level. No effect related to the report versus no-report conditions was obtained ($F = .006$, $df = 1/72$, $p < .93$). The within-subject effect of trials reached significance ($F = 437.98$, $df = 5/360$, $p < .001$). All interactions were nonsignificant.

The correlations of total number of correct responses, with the sums of each of the self-report scales over all six trials, are shown in Table I. The 13 affects were separated into positive and negative categories along the general lines previously established by Tomkins.[11] Grouping all subjects together, the positive affects generally correlated positively with performance, while the negative affects correlated negatively. Separate correlations for the high-, medium-, and low-anxiety groups were performed. The low- and medium-anxiety groups showed considerable similarity in the pattern of correlations, while the high-anxiety group differed in both sign and magnitude from the other two groups. For the low- and medium-anxiety groups, the negative self-report scales generally correlated negatively with performance, and the positive scales correlated

---

[11] S. S. Tomkins, *Affect, Imagery, Consciousness*, 1962, 337.

TABLE I

CORRELATIONS OF POSITIVE AND NEGATIVE SELF-REPORTED AFFECTS WITH TOTAL CORRECT RESPONSES AS A FUNCTION OF SUBJECT GROUPING

	Negative affects						
	Bored	Annoyed	Unpleasant	Anxious	Tense	Discouraged	Ashamed
All *Ss*	−.281	−.493***	−.504***	−.036	.065	−.185	−.168
HA *Ss*	−.355	−.277	−.257	.335	.412	.186	.148
MA *Ss*	−.281	−.592**	−.614**	−.194	−.344	−.268	−.240
LA *Ss*	−.275	−.478	−.521*	−.037	.111	−.253	−.487
	Positive affects						
	Excited	Interested	Calm	Satisfied	Relaxed	Pleasant	
All *Ss*	.206	.347**	.116	.139	−.040	.167	
HA *Ss*	.653***	.738***	−.530*	−.347	−.447	−.266	
MA *Ss*	.072	.017	.500*	.200	.271	.265	
LA *Ss*	.065	.325	.231	.517*	.001	.346	

* $p < .05$.
** $p < .02$.
*** $p < .01$.

positively with performance. However, for the high-anxiety group, the negative affects *anxious, tense, discouraged,* and *ashamed* correlated positively with performance, while the positive affects *calm, satisfied, relaxed,* and *pleasant* correlated negatively with performance. The positive affects *excited* and *interested* correlated significantly and positively with performance for high-anxiety subjects.

Averages for the 13 self-reported affects were obtained at six points in the experiment. Generally, for all subjects grouped together, all affects stabilized by the second trial, so that there were no significant variations thereafter. For example, the only three affects significantly related to performance for all subjects (*annoyance, interest,* and *unpleasantness*) each showed significant changes between the first and second trials, and showed no differences thereafter. Between the first and second trials, related t tests showed that *annoyance* increased significantly ($t = 3.44$, $p < .005$), *unpleasantness* increased significantly ($t = 2.73$, $p < .005$), and *interest* decreased significantly ($t = 2.14$, $p < .025$). No significant differences for *annoyance, unpleasantness,* and *interest* occurred after the second trial.

Because variance in affect level was minimal after the second trial, the correlations of affects for all subjects, with number of responses correct on each trial, were inspected at all six trials in the experiment. (The reader will recall that affects were rated *prior* to each performance trial.) It was only on the third trial that the same three affects which correlated significantly with performance summed over all six trials (*annoyance, unpleasantness,* and *interest*) correlated significantly with performance on that trial. After the third trial, the correlations of these three affects and performance, trial by trial, remained constant and continued to be significant.

In order to ascertain whether feedback manipulations were associated with affective changes, comparisons were made of the positive and negative self-report scales for each feedback condition. Summing across positive affects, there were no differences across feedback conditions, with the positive being nonsignificantly different from the average feedback ($t = .62$, $p < .90$), and the negative being nonsignificantly different from the positive feedback ($t = .36$, $p < .80$). There was a nonsignificant trend going from the positive- to the average- and negative-feedback conditions for the negative self-reported affects to increase. Summing across negative affects there were no significant differences between the positive and average feedback ($t = .34$, $p < .80$), no significant differences between the average and negative feedback ($t = .95$, $p < .30$), and no significant differences between the negative and positive feedback ($t = 1.37, p < .15$).

Summing over positive and negative self-reported affects, *t* tests showed no significant differences as a function of STAI anxiety level. Thus, the significant main effect of anxiety was not mediated through either positive or negative self-reported affects.

DISCUSSION

According to Spence and Spence, the most widely accepted position among anxiety theorists is that high-anxiety subjects react more to stress than low-anxiety subjects, and that the competing responses evoked by this emotionality result in the high-anxiety subjects being inferior in performance on a complex learning task.[12] Using a situational index of anxiety such as the STAI employed in this study, this prediction is upheld: higher anxiety was associated with poorer performance. Using the same task, results in our laboratory have consistently shown that high-anxiety subjects are inferior in performance.[13] The task utilized in this study was employed because of its correlation with performance in the classroom, so the results have implications in this area.

While the effect of feedback was not significant by analysis of variance, there was a consistent trend for performance to improve going from the negative to average to positive conditions. That the trend was slight is indicated by the .13 level associated with feedback in the analysis of variance. Although the differences were slight (see Figure 1), they were highly consistent in that for each anxiety level and report condition, the more positive feedback was associated with better performance in 12 possible comparisons. Thus, the nonparametric sign test was employed to ascertain the possibility of achieving results this consistent in nature, and this trend was significant at the .001 level. While differences due to feedback were small, their consistency suggests that manipulation of feedback by offering the student a low, average, or high level of comparison with his peers might profit from further research. An especially useful finding, for actual application to the classroom setting, would be to ascertain what level of comparison optimally 'motivates' a particular type of student. Such an experiment is under progress in our laboratory.

---

[12] Spence and Spence (see n. 5 above).

[13] Katahn and Branham (see n. 9); William J. Ray, Effects of test anxiety on acquisition, retention, and generalization of meaningful verbal material in a classroom situation, M.A. thesis, Vanderbilt University, 1969.

There were no differences in self-reported affect as a function of feedback or anxiety. Self-reported affect did significantly relate to performance in many cases, but this affect was associated neither with the feedback manipulations nor with anxiety groups. Feedback tended to be associated with different levels of performances, while anxiety was strongly related to performance differences. However, affect as indicated by self-report was not the mediating agent by which differences were obtained. Thus, self-reported affect, feedback, and anxiety all appear to influence performance, but as far as measured in this study, they do so independently of each other. These failures of the feedback conditions and anxiety grouping to significantly influence emotional reactions point out the importance of verifying the effects of experimental manipulations when such manipulations are simply assumed, on some a priori basis, to have an effect on the subject's internal state.

Results indicated that performance did not differ whether a subject did or did not take self-report scales throughout the task. Generally for all subjects in the report conditions, higher self-report ratings on affects which are conceived as positive in nature were associated with better performance. Conversely, higher self-report ratings on affects typically conceived as negative were associated with poorer performance. Additionally, the direction of causality relating affects and performance can be determined in the present study. In the current study, the primary change in affect occurred by the second trial, as indicated by the significant increase in both positive and negative affects going from the first to second trials, with relatively little change in affect level thereafter. For all subjects, affect did not significantly relate to performance on either the first or the second trial, but affect did in many cases significantly relate to performance when the total numbers of correct responses were correlated with summed-affect scores over six trials. Further, trial-by-trial analyses of affects and performance showed that the only significant *change* in affect levels occurred between the first and second trials. Thereafter, affect levels remained fairly constant. Affect and performance did not begin to correlate until the third trial. Therefore, changes in affect occurred before the establishment of any significant affect-performance correlations. This finding is consistent with other research in our laboratory. Using continuously gathered self-report data during serial learning, Katahn, Snyder, and Durlak found that the main performance decrement followed the occurrence of increased negative affect, and not vice versa.[14]

One point of interest, which does not lend itself to ready interpretation, is the difference in pattern of correlations between affect and performance for high-anxiety subjects as compared with low- and medium-anxiety subjects. While high-anxiety subjects were significantly inferior to the subjects in the other two groups, they performed better (within their group) as some negative affects increased, whereas low- and medium-anxiety subjects performed better as positive affects increased. Looking only at the significant correlations of self-reported affect and performance, the composites of optimal affective states for the groups were as follows: (*a*) high-anxiety subjects performed best when they reported high on *excitement* and *interest*, and low on *calmness;* (*b*) medium-anxiety subjects performed best when they reported high on *calmness* and low on *annoyance* and *unpleasantness;* (*c*) low-anxiety subjects performed best when they reported high on *satisfaction* and low on *unpleasantness*. Because *interest* seemed to be a critical self-report scale for high-anxiety subjects, the intercorrelations of the other scales with *interest* were examined. For the high-anxiety group, *interest* correlated positively with the majority of negative affects, while for the medium- and low-anxiety groups, *interest* correlated positively with a majority of the positive affects. No ready explanation presents itself to account for this difference in affect-scale relationships, among themselves and to performance, for groups varying in test anxiety. However, another study relating a full range of affect scores to performance is now in progress. A replication of the pattern of different affect-performance relationships as a function of test anxiety would present a considerable theoretical challenge.

It is noteworthy that state anxiety, as measured by the STAI, was significantly related to performance, while scores on the nine-point anxiety scale collected during performance were not, considered over all subjects. Thus the relationship of affect to performance may well depend on the type of scale used to measure it, and the strange pattern of affect-scale correlations within different STAI anxiety groups may be purely an artifact of the measuring instruments.

---

[14] Katahn, Snyder, and Durlak (see n. 8).

The STAI scores accounted for 31% of the variance. The anxiety self-report scale accounted for from 1% to 11% of the variance (depending on anxiety group), and the other affect scales related to from 1% to 50% of the variance. If the goal is the prediction of performance from variation in affect, it is suggested that future research utilize the other affects in addition to anxiety, and that measures of affect be taken during the actual testing situation. Further theoretical refinements may be achieved should consistent results relating the whole range of affects to performance be obtained.

GENERALIZATION OF POSITION ASSOCIATION IN ROTE SERIAL LEARNING

MUKUL K. DEY, *McMaster University, Canada*

Abstract. Thirty-six subjects practiced two serial lists of identical length in haphazard alternation. At each position in a list, a maximum number of intruding responses arrived from the same position in the other list, and the frequency of intrusion from a nonidentical position in the other list declined with increased positional disparity in either direction. The reliability of this gradientlike phenomenon, which could not be verified by the method of trend analysis because of insufficient data, is partially indicated by the occurrence of interlist intrusions at a significantly higher rate between identical than between nonidentical positions. These results are interpreted as evidence in support of the hypothesis that associations not only develop between the items of a rote series and their serial positions but also generalize between positions in inverse proportion to the intervening distance.

Rote serial learning is traditionally equated with the formation of a chain of associations between successive list items. This traditional view is now losing ground in the wake of the revival of a view which accords the functional stimulus role to the position an item holds in the series.[1] The large mass of experimental data recently generated by the revived interest in the position-association hypothesis is not unequivocal, but the evidence contradicting the hypothesis is outweighed by that supporting it.

For example, first, the hypothesis of association with serial position predicts that since serial-position cues are entirely absent in the varied-order paired-associate procedure, indeterminate transfer should result when a serial list is followed by a paired-associates list composed of the same items, regardless of whether the formation of pairs is or is not consistent with the items' order in the serial list. This prediction is supported by the literature, although negative evidence has been reported.[2]

Received for publication December 16, 1969.

[1] R. S. Woodworth and A. T. Poffenberger, *Textbook of Experimental Psychology*, 1920 (mimeo ed.).

[2] For the positive evidence, see C. C. Erickson, R. D. Ingram, and R. K. Young, Paired-associate learning as a function of rate of presentation and prior serial learning, this JOURNAL, 76, 1963, 458–463; A. R. Jensen and W. D. Rohwer, What is learned in serial learning? *J. verb. Learn. verb. Behav.*, 4, 1965, 62–72; R. K. Young, A comparison of two methods of learning serial association, this JOURNAL, 72, 1959, 554–559; R. K. Young, The stimulus in serial verbal learning, this JOURNAL, 74, 1961, 517–528; R. K. Young, Tests of three hypotheses about the effective stimulus in serial learning, *J. exp. Psy-*

Second, the learning of a serial list becomes considerably more difficult if the point in the series where list presentation begins is always changed between trials than if it is unchanged.[3] Assuming that serial position provides a wider and, therefore, more effective spectrum of discriminative cues than the items themselves normally do, the above phenomenon can be interpreted as solid evidence in favor of the position hypothesis.

Third, greater retroactive inhibition is reported to occur in serial learning when the starting point is unchanged than when it is changed through the trials,[4] and in paired-associates learning when the order of pairs is constant than when it is varied.[5] Since the conditions compared in the experiments cited differ only with regard to the presence of serial organization, the additional interference effect is presumably due to no other process than the association of rival responses with common serial-position cues.

Fourth, the clearest evidence comes from the studies of transfer of training from an original serial list to a derived serial list. Young and Keppel and Saufley have shown that in second-list learning the items repeated from the first list at identical serial

---

chol., 63, 1962, 307–313; R. K. Young and M. Casey, Transfer from serial to paired-associate learning, *J. exp. Psychol.*, 67, 1964, 594–595. For the negative evidence, see R. S. Heaps, W. A. Greene, and C. D. Cheney, Transfer from serial to paired-associate learning with two paired-associate rates, *J. verb. Learn. verb. Behav.*, 7, 1968, 840–841; L. M. Horowitz and C. Izawa, Comparison of serial and paired-associate learning, *J. exp. Psychol.*, 65. 1963, 352–361; L. Postman and K. Stark, Studies of learning to learn; IV, Transfer from serial to paired-associate learning, *J. verb. Learn. verb. Behav.*, 6, 1967, 339–353; T. J. Shuell and G. Keppel, A further test of the chaining hypothesis of serial learning, *J. verb. Learn. verb. Behav.*, 6, 1967, 439–445.

[3] R. E. Bowman and W. R. Thurlow, Determination of the effects of position in serial learning, this JOURNAL, 76, 1963, 736–745; S. M. Ebenholtz, Serial learning: Position learning and sequential associations, *J. exp. Psychol.*, 66, 1966, 353–362; G. Keppel, Retroactive inhibition of serial lists as a function of the presence or absence of positional cues, *J. verb. Learn. verb. Behav.*, 3, 1964, 511–517; Shuell and Keppel (see the last reference in n. 2 above); W. A. Winnick and R. L. Dornbush, Role of positional cues in serial rote learning, *J. exp. Psychol.*, 66, 1963, 419–421; B. Zavortink and G. Keppel, Retroactive inhibition of interitem associations for serial lists with constant or varied starting positions, *J. verb. Learn. verb. Behav.*, 7, 1968, 771–775.

[4] Keppel (see n. 3 above).

[5] J. A. McGeoch and F. McKinney, Studies of retroactive inhibition: The influence of relative order of presentation of original and interpolated paired associates, *J. exp. Psychol.*, 20, 1937, 60–83; J. A. McGeoch and B. J. Underwood, Test of the two-factor theory of retroactive inhibition, *J. exp. Psychol.*, 32, 1943, 1–16.

positions are learned faster than those repeated at disparate serial positions.[6] Further, Ebenholtz has demonstrated quicker learning of a derived list containing some of the first-list items at identical serial positions in comparison with another derived list in which the same items are located a few steps away from their original serial positions.[7]

As the position hypothesis thus appears to have enough support to merit serious consideration, a pertinent question is whether the response associated with a certain serial position is apt to be generalized to other serial positions and whether the degree of such generalization is a monotonic negative function of positional disparity. This question, though not formulated in exactly the same way, was subjected to evaluation by Schulz and by Hicks, Hakes, and Young.[8] Having their subjects name the serial positions at the items' random presentation after the completion of serial learning, these investigators demonstrated that the distribution of errors in naming the serial positions closely conforms to the bow-shaped serial-position gradient. In a strict sense, this phenomenon is no evidence of serial-position generalization, though it does suggest that confusing the serial positions may be either a cause of the serial-position gradient or an effect of the same process which causes that gradient. The study reported in this paper arranged for a direct inquiry into the possible existence of serial-position generalization. The subjects were asked to learn two separate lists of identical length through a practice sequence in which the two lists were unsystematically alternated. The plan was to analyze the transfer of overt response between lists—transfer which the procedure outlined was expected to generate in a sufficient quantity. If the results revealed an inverse relationship between the interlist transfer of specific responses and the longitudinal distance between serial positions, such a finding would not only demonstrate serial-position generalization but also provide further strength for the position hypothesis.

---

[6] Young (see his 1962 article in n. 2 above); G. Keppel and W. H. Saufley, Jr., Serial position as a stimulus in serial learning, *J. verb. Learn. verb. Behav.*, 3, 1964, 335–343.

[7] Ebenholtz (see n. 3 above).

[8] R. W. Schulz, Generalization of serial position in rote serial learning, *J. exp. Psychol.*, 49, 1955, 267–272; R. Y. Hicks, D. T. Hakes, and R. K. Young, Supplementary report: Generalization of serial position in rote learning, *J. exp. Psychol.*, 78, 1968, 536–538.

METHOD

Subjects. Thirty-six men, part-time undergraduate students from several evening classes in psychology, volunteered as *S*s. Their mean age was 26.7 years in the approximate range from 19 to 36.5 years. All had previously participated as observers in experiments on paired-associates as well as serial learning.

Apparatus and material. The task was presented by means of a compact electromagnetic unit devised by Ranschburg in 1901.[9] This unit rhythmically turns a cardboard disc in steps, thereby exposing, one by one, the items printed in the narrow sectors of the disc. A recycling timer regulated the exposure rate.

The two six-item lists contained easily pronounceable nonsense syllables chosen from the medium range of association values in Glaze's inventory.[10] In constructing the lists, these rules were observed: (*a*) with the exception of the first syllable, no two syllables within either list had the same vowel or any common consonant; (*b*) alphabetical sequences of letters within a syllable and from one syllable to the next in the series were avoided as far as possible; (*c*) the syllables with the same vowel or any common consonant in the two lists were never placed at identical ordinal positions. Further, instead of testing all 36 *S*s with the same pair of lists, four separate pairs of lists were prepared, each for use with 9 *S*s. In each list, the first syllable was printed in red boldface type and the remaining syllables were printed in black boldface type. It was assumed that by its distinctive color, the first syllable enabled *S* to perceive the beginning and the end of each list from the start of practice.

Procedure. Practice was administered by the anticipation method. The two lists were successively exposed once; then, at the presentation of each item, *S* was required to anticipate by spelling out the item coming next. The time available to *S* for anticipation was roughly the same as the 4-sec. exposure duration of each item. In the instructions, *S* was informed that the red item was the starting point in each list, and further, in order to place emphasis on overt responding, *S* was asked to guess in the event of uncertainty.

The items of each list always appeared in the same sequence, but the two lists were alternated in the order 1221211212. This cycle was repeated if *S* needed more practice to reach the mastery criterion of first correct elicitation of all ten responses in the successive presentation of the two lists. With the absence of any interval between list presentations, the task was continuous. The last item in each list had no intended stimulus function, and therefore, though *S* frequently reacted to it, especially at the early stage (being under instruction to anticipate whenever possible and having been provided no information about the length of each list), the *E*'s measurement was concerned with *S*'s reactions to the first five items only. A correct anticipation was considered a failure if the response was initiated too late to be entirely within the anticipation interval. All overt responses and omissions were recorded by *E* on a mimeographed data sheet.

---

[9] Though antiquated, this device delivers rote-learning materials as efficiently as any other exposure apparatus of which the writer knows.

[10] J. A. Glaze, The association value of nonsense syllables, *J. genet. Psychol.*, 35, 1928, 255–267.

RESULTS

The data consist of the reactions elicited by the first five items in each list. On the average, each subject gave 88.6 reactions per list. Among these, there were 47.03 (53.08%) correct responses and 41.57 (46.92%) errors. The errors were composed of 10.2 (11.51%) omissions, 4.44 (5.01%) complete intralist intrusions, 10.99 (12.41%) complete interlist intrusions, and 15.94 (17.99%) miscellaneous responses. The last category was mostly made up of incomplete or partial intrusions and reactions citing the first items.

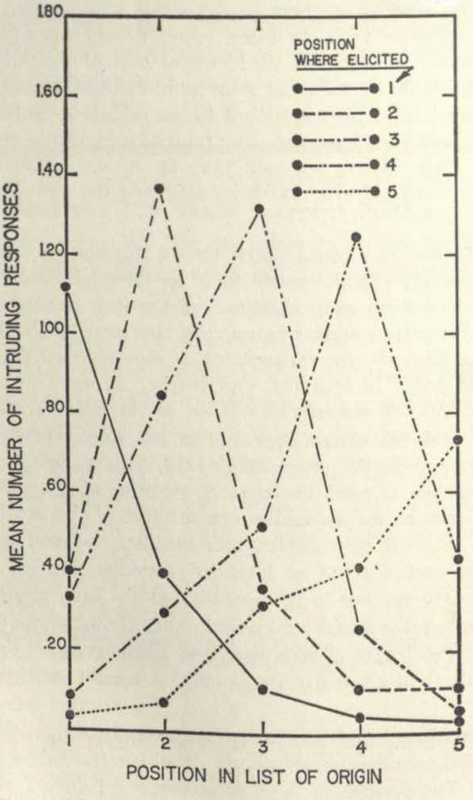


FIG. 1. A breakdown of the interlist intrusions

Interlist intrusions. Out of the average 10.99 interlist intrusions per list, 5.77 were from identical or coordinate serial positions and 5.22 were between nonidentical or disparate serial positions. If it is true that the serial-position cues operate as functional stimuli, then there should have been considerable overt transfer of response between identical as compared with nonidentical serial positions in the two lists. Since the scope of measuring overt transfer of response of the former category was four times as great as the scope of measuring overt transfer of response of the latter category in this experiment, a proper evaluation of the above hypothesis required that the number of intrusions between identical serial positions be multiplied four times. A test for the significance of difference between the so-weighted mean number of intrusions between identical serial positions (23.08) and the mean number of intrusions between nonidentical serial positions (5.22) yielded a t ratio of 11.39, which, for 39 *df*, was significant beyond the .01 level. Similar analyses were made for the five positions separately. The weighted mean frequencies of intrusion between coordinate positions were, respectively, 4.44, 5.44, 5.28, 5, and 2.92, and the mean frequencies of intrusion between disparate positions were, respectively, .58, .97, 1.5, 1.33, and .84 for the first through fifth positions. Tests for the reliability of difference between the above five pairs of means produced the following values of t, all of which were significant beyond the .05 level: 6.92, 4.49, 5.2, 4.34, and 4.76.

Figure 1 provides a breakdown of the interlist intrusions in terms of the intruding items' positions in the list of origin. The five curves portray the mean frequencies with which the item at each of the five positions in one list elicited the responses located at the five positions in the other list. All five curves vividly suggest that a maximum number of intruding responses arrived from the same serial position in the other list and that the frequency of intrusion from a nonidentical serial position in the other list systematically declined as positional disparity increased in either forward or backward direction. Unfortunately, being of insufficient volume, the data do not lend themselves to the statistical procedure by which the reliability of these gradientlike trends could be ascertained.

Intralist intrusions. In order to determine whether the interlist and intralist intrusions were elicited at different rates, a t-test comparison was made of the mean frequencies of interlist intrusions between disparate serial positions (5.22) and of intralist intrusions (4.44). The results reveal lack of significance ($t = 1.59$, $p > .1$). The mean intralist intrusions from the anterior and posterior directions were, respectively, 3.88 and .56. A t ratio of 7.02 ($p < .01$) attested the reliability of the difference between these means. Similar instances of preponderance of intralist intrusions from the anterior direction over those from the posterior direction in serial verbal learning were earlier reported by Bugelski and by Slamecka.[11]

DISCUSSION

The rate of interlist overt transfer of response between coordinate positions was significantly higher than that between disparate positions. Further, interlist overt transfer appeared to have a gradient-like relationship with degree of positional disparity. Though insufficient data precluded tests for statistical significance of the latter finding, its consistent and systematic character perhaps vouches for its dependability. Part of the interlist intrusions were admittedly due to formal similarity between the items, but as this factor presumably wields a uniform influence, it cannot be held responsible for the observed variations. Therefore, the above findings can be best interpreted as the reflection of a tendency for the serial positions to generalize in conformity with the law of stimulus-generalization gradient.

The process underlying the operation of the serial positions as functional stimuli is subject to speculation. A simple but testworthy hypothesis is that an awareness of the ordinal-position numbers of the items develops either spontaneously or through a conscious search for mediating cues and that associations are then developed between these numbers and the corresponding responses. This hypothesis was tested by Jensen and Blank in an experiment on the serial learning of forms.[12] Though their experiment produced negative evidence by showing that continuous briefing of the subject on the ordinal-position numbers of items did not facilitate learning, some positive evidence was furnished by the reports of positive transfer from a serial task to a paired-associates task derived by pairing the same verbal items with their ordinal-position numbers[13] or with the points in a spatial dimension that were coordinate with the ordinal positions previously occupied by the verbal items in the serial task.[14]

---

[11] B. R. Bugelski, A remote association explanation of the relative difficulty of learning nonsense syllables in a serial list, *J. exp. Psychol.*, 40, 1950, 336–348; N. Slamecka, An inquiry into the doctrine of remote association, *Psychol. Rev.*, 71, 1964, 61–76.

[12] A. R. Jensen and S. S. Blank, Association with serial position in serial rote learning, *Canad. J. Psychol.*, 16, 1962, 60–63.

Alternatively, the serial-position cue may be interpreted as the time interval elapsing between the start of the series and the point at which the appropriate response is located. Such an interpretation appears plausible in the light of Czehura's experiment on instrumental conditioning, which demonstrated that the time interval between very brief events occurring in rapid succession may function as a stumulus and that the response conditioned to one time interval may undergo generalization to another time interval in inverse proportion to their difference.[15] Further, it follows from this interpretation that, with the first item in a serial list being the initiating point of every time interval that takes on stimulus function, the factors promoting distinguishability should be important variables in serial organization. By demonstrating the steepening of the serial-position gradient by one such factor—the rest pause between list repetitions—the experiments of Glanzer and Peters and of Lippman and Denny add credibility to this alternative interpretation of the serial-position cue.[16]

---

[13] W. A. Winnick and R. L. Dornbush, Ordinal position in serial learning, *J. exp. Psychol.*, 78, 1968, 536–538.

[14] S. M. Ebenholtz, Position mediated transfer between serial learning and a spatial discrimination task, *J. exp. Psychol.*, 65, 1963, 603–608.

[15] W. S. Czehura, The generalization of temporal stimulus patterns on the time continuum, *J. comp. physiol. Psychol.*, 36, 1948, 79–90.

[16] M. Glanzer and S. C. Peters, Re-examination of the serial position effect, *J. exp. Psychol.*, 64, 1962, 258–266; L. G. Lippman and M. R. Denny, Serial position effect as a function of intertrial interval, *J. verb. Learn. verb. Behav.*, 3, 1964, 496–501.

UNLEARNING OF CONTEXT-ITEM AND SPECIFIC ASSOCIATIONS IN THE W_1–R_1, W_2–R_2 PARADIGM

N. Jack Kanak and C. Dwayne Curtis, *University of Oklahoma*

Abstract. Unlearning of List 1 contextual associations was studied in the verbal-discrimination nonspecific-transfer paradigm (W_1–R_1, W_2–R_2). Bidirectional modified free-recall data revealed significant experimental unlearning of context-W (wrong item) and context-R (right item) associations, compared with both a control group that recalled immediately after List 1 practice and a 'normal forgetting' control group that recalled List 1 after working on a neutral interpolated task. In addition, an associative-matching task indicated that the experimental group also experienced unlearning of specific, incidentally learned, W-R/R-W List 1 associations. These results represent an extension to verbal-discrimination learning of several recent paired-associate studies that demonstrated unlearning of specific as well as context-item associations in the A–B, C–D paradigm.

McGovern presented evidence in a study of the A–B, C–D paired-associate nonspecific-transfer paradigm that List 1 contextual associations formed between the learning context and the response items undergo extinction, or unlearning, during C–D learning.[1] Specific forward (A–B) and backward (B–A) associations, however, were shown to remain near criterion strength. The unlearning of context-response associations was measured in a modified free-recall task. Recall of specific associations was measured by an associative-matching task, which is assumed to eliminate the effects of item unlearning. The extinction of context-item associations was attributed to the fact that the context-response associations of the two successive lists enter into the well-documented A–B, A–C competition-unlearning relationship.

Houston noted that the absence in the McGovern study of any evidence for the unlearning of specific List 1 associations could be an artifact of either (*a*) the overlearning of List 1 associations or (*b*) the failure to utilize a control group required to perform an associative-matching task after only single-list learning.[2] In

---

Received for publication October 31, 1969. A paper based on the study was presented at the annual meeting of the Midwestern Psychological Association, Chicago, 1969.

[1] J. B. McGovern, Extinction of associations in four transfer paradigms, *Psychol. Monogr.*, 78, 1964 (No. 16, Whole No. 593).

[2] J. P. Houston, Unlearning of specific associations in the A-B, C-D paradigm, *J. exp. Psychol.*, 74, 1967, 254–285.

his own study, Houston employed a 50% List 1 learning criterion and included the baseline-recall control group for both the modified free-recall and the associative-matching measures. Consistent with McGovern's results, the recall data did reveal the unlearning of context-response associations. More importantly, in contrast to McGovern, Houston's matching task indicated the unlearning of specific List 1 associations as well. Garskof, and Garskof and Bryan, using 100% and 50% original-learning criterions respectively, have replicated Houston's results.[3]

Abra, however, reports an experiment that failed to reveal extinction of either context-item or specific List 1 associations following C–D learning.[4] The discrepancies between his results and those of Houston and Garskof may be attributable to a potentially strong list-differentiation factor in Abra's materials. Abra utilized low-frequency nouns as List 1 stimuli, but he used complex perceptual figures as stimuli in List 2 in order to achieve minimal interlist stimulus similarity. It is conceivable that this procedure permitted the maintenance of independent List 1–List 2 context-item and context–specific-association systems based on stimulus form-class as a cue to interlist differentiation. The mechanism implicated in this suggestion may be analogous to the mechanism involved in the decrement of retroactive inhibition that has been observed in the A–B, A–C paradigm when the interlist form-class of the response terms is varied.[5]

The present study extends the investigation of the unlearning of specific and context-item associations to the verbal-discrimination analogue of the A–B, C–D paradigm. The comparable paradigm is designated W_1–R_1, W_2–R_2, a notation indicating that the List 2 wrong (W) and right (R) items are unrelated to List 1 W and R items. Bidirectional W–R and R–W associative learning and the learning of W and R items as responses have been demonstrated to be incidental learning components of the verbal-discrimination task.[6] In addition, recent transfer studies have shown that these incidentally learned components activate processes in verbal-discrimination transfer and retroaction research that are comparable to the operation of associative mechanisms in paired-associate transfer tasks.[7] The demonstration of unlearning of context-item and specific associations in the verbal-discrimination nonspecific-transfer paradigm would thus represent an expansion of the growing number of phenomena and processes which have been found to exist in common between verbal-discrimination and paired-associate tasks.

---

[3] B. E. Garskof, Unlearning as a function of degree of interpolated learning and method of testing in the A-B, A-C and A-B, C-D paradigms, *J. exp. Psychol.*, 76, 1968, 579–583; B. E. Garskof and T. M. Bryan, Unlearning as a function of degree of original learning and retention test, *Psychon. Sci.*, 6, 1966, 391–392.

[4] J. C. Abra, Supplementary reports: Time changes in the strength of extinguished context and specific associations, *J. exp. Psychol.*, 77, 1968, 684–686.

[5] L. R. Goulet and R. N. Bone, Interlist form-class similarity and interference effects, *J. verb. Learn. verb. Behav.*, 7, 1968, 92–97; L. Postman, G. Keppel, and K. Stark, Unlearning as a function of the relationship between successive response classes, *J. exp. Psychol.*, 69, 1965, 111–118.

In the W_1–R_1, W_2–R_2 paradigm, associations formed between the context and W items and between the context and the R items of List 1 both enter into an A–B, A–C competition-unlearning relationship with their context-W and context-R counterparts of List 2. Resultant unlearning of both types of List 1 context-item associations was therefore expected to occur during W_2–R_2 learning and was expected to be revealed during bidirectional modified free recall. The unlearning of incidentally learned W–R/R–W specific associations was tested in an associative-matching task. Specific associations in the W_1–R_1, W_2–R_2 paradigm conceivably are subject to the same unlearning processes as the intentional associations of the paired-associate nonspecific-transfer paradigm, though apparently little agreement exists concerning these processes in paired-associate learning.[8]

METHOD

Subjects and experimental design. Forty-eight students from introductory psychology courses at the University of Oklahoma volunteered to serve as *S*s. None had served before in verbal-learning experiments. The *S*s were unsystematically assigned to the six treatment conditions (3 groups × 2 list variations) in order of their appearance at the laboratory, within the restriction that *n* cells be filled before the *n*+1 cell was assigned.

---

[6] D. H. Kausler and R. J. Sardello, Item recall in verbal-discrimination learning as related to pronunciation and degree of practice, *Psychon. Sci.*, 7, 1967, 285–286; R. J. Sardello and D. H. Kausler, Associative recall in verbal-discrimination learning as related to pronunciation and degree of practice, *Psychon. Sci.*, 8, 1968, 251–252.

[7] N. J. Kanak and M. F. Dean, Transfer mechanisms in verbal discrimination, *J. exp. Psychol.*, 79, 1966, 300–307; D. H. Kausler, F. E. Fulkerson, and A. J. Eschenbrenner, Unlearning of List 1 right items in verbal-discrimination transfer, *J. exp. Psychol.*, 75, 1967, 379–385.

[8] Garskof (see n. 3 above).

One control group (Group IM) was given the modified free-recall and associative-matching tasks immediately after List 1 learning. The experimental group (Group ITT) recalled List 1 after practice on the interpolated transfer task (the W_2–R_2 list). The third group (Group NIT) recalled List 1 after performing a neutral interpolated task (the Meier Art Judgment Test) for a time period equal to ten List 2 trials for Group ITT. Group NIT therefore served as a control for 'normal forgetting' over time.[9]

List construction. A pool of 60 familiar words was selected from normative materials[10] with the primary criterion for selection being that associative relations between all words be minimized. Words of less than three or more than eight letters, compound words, and words with associates which followed natural language habits (e.g., 'very much') were excluded. From the initial pool of 60 words, two subpools of 15 words each were randomly selected to serve as W and R items. Random pairings of the W and R items were then created, although idiosyncratic pairings were avoided. This set of 15 pairs was designated List 1A. List 1B was formed by so re-pairing the items of List 1A that the two lists differed only in the specific word-pairings. Lists 2A and 2B for the experimental group were formed in the same way from the remaining 30 words. The two sets of list pairings were utilized to permit greater generality over particular word pairings.

Four random serial orders of pairs and four orders of right-left spatial positions of the W or R items were utilized in each set to eliminate serial-position and spatial-position learning effects. In a given random order, the R item occurred either seven or eight times on each side, within the restriction that successive runs of a spatial position be limited to four items. In addition, across the four serial orders a given R item occurred on the left or right an equal number of times.

Procedure. The lists were presented by the anticipation method on a Lafayette memory drum. During the anticipation exposure, both items (W and R) appeared in the window of the memory drum and *S* pronounced aloud the item selected as correct. During the feedback exposure, both items again appeared contiguously but with the R item underlined. Temporal conditions were set at a 2:2 sec. exposure rate with a 4-sec. intertrial interval. Standard verbal-discrimination instructions were given along with the information that neither serial-order nor spatial-position cues would be helpful. The *S*s were not informed in advance that recall measures would be taken or that a second task might follow. Lists 1A and 1B were counterbalanced across *S*s in all three treatment conditions.

List 1 learning was to a criterion of one errorless trial. Group ITT *S*s began ten trials of practice on List 2 approximately 1 min. after reaching the List 1 criterion. Group NIT *S*s practiced on a neutral interpolated task for 10 min., a time period equivalent to the ten List 2 trials for Group ITT. Immediately following List 1 practice for Group IM, or following the interpolated tasks for Groups ITT and NIT, a counterbalanced administration of bidirectional modified free recall of List 1 items was introduced. Half of the *Ss* of each group were asked first to recall R items to given W items (typed in a new random order on a sheet of paper) and then to recall W items to given R items. The remaining *Ss* received the counterbalanced, reverse order of administration. The instructions emphasized the recall of List 1 items for Group ITT. Unlimited time was allowed for recall and all *Ss* were instructed to guess if uncertain. Following recall, *Ss* of all groups were given a sheet of paper listing all the W and R items of List 1 in new random orders in separate columns. The *Ss* were instructed to match the items as they occurred in the list. Again, unlimited time was allowed for the associative-matching task with instructions for guessing.

---

[9] J. M. Barnes and B. J. Underwood, "Fate" of first-list associations in transfer theory, *J. exp. Psychol.*, 58, 1959, 97–105.

[10] D. S. Palermo and J. J. Jenkins, *Word Association Norms: Grade School through College*, 1964.

RESULTS AND DISCUSSION

Acquisition. Lists 1A and 1B were found comparable both on mean trials to criterion and on means for total errors (Fs < 1). The two sets of lists were therefore combined within each group for the overall between-groups tests for differences in List 1 learning. The mean trials to criterion ranged from 7.44 to 7.81, while the comparable means for errors ranged from 26.44 to 29.38. Since both measures resulted in Fs $(2, 45) < 1$, the groups were considered adequately equated in degree and difficulty of List 1 learning. In addition, the modified-free-recall and associative-matching scores were found to be entirely independent of sets of lists, permitting the omission of the list-set factor in all the statistical analyses reported below.

Unlearning of context-item associations. Correct responses in modified free recall were examined in a 3 × 2 × 2 factorial analysis of variance, with groups and order of recall as between-subject measures, and direction of recall as a repeated measure. The overall difference between groups in total recall was the only significant effect [$F\,(2, 42) = 5.53$, $p < .01$]. The means and standard deviations, respectively, were 5.75 and 5.54 (Group ITT), 12.44 and 7.28 (Group NIT), and 11.62 and 6.50 (Group IM). Duncan's multiple-range test indicated a significant loss of W and R item availability for Group ITT relative to both Groups IM and NIT (ps $< .01$). The small difference between the latter two groups did not approach significance ($p > .10$). The unlearning of context–W and context–R associations in Group ITT is consistent with the unlearning of context-item associations observed by several investigators following List 2 C-D learning in the paired-associates task. Context-item unlearning in the verbal-discrimination task may similarly be explained as the result of the context-item associations of two successive lists entering into an A–B, A–C competition-unlearning relationship.

Unlearning of specific associations. The analysis of variance of correct responses on the associative-matching task revealed a significant between-group effect [F (2, 45) = 3.46, $p < .05$]. The means and standard deviations, respectively, were 4.19 and 3.82 (Group ITT), 7.19 and 3.90 (Group NIT), and 7.31 and 4.06 (Group IM). Duncan's test indicated that significant unlearning of specific associations occurred in Group ITT relative to both control groups (both ps $< .05$). This result seemingly represents an extension to verbal-discrimination learning of the phenomena of associative unlearning that was documented in the paired-associate studies of Houston and Garskof.

The procedure of utilizing the same subjects for both the modified-free-recall and the associative-matching tasks may, however, represent a potential source of confounding mitigating the above conclusion. For example, Keppel has shown, in a single-list verbal-discrimination learning situation, that associative-matching scores tended to be significantly depressed when the same subjects first experience modified free recall.[11] Keppel suggests that errors in prior recall may have been responsible for this depression of matching scores. If this effect were generalizable to the ITT group of the present study and, more importantly, if the effect were greater for the experimental group than for the control groups, the observed retroactive inhibition of specific associations could be artifactual.

To evaluate this possibility, two additional analyses were conducted. The number of incorrect responses (items recalled but paired incorrectly) in recall were compared in a 3 (groups) × 2 (incorrect R- versus W- item recall) factorial analysis of variance, with the latter factor a repeated measure. More R items ($\bar{X} = 3.25$) were incorrectly paired than W items ($\bar{X} = 2.39$) [F (1, 45) = 6.27, $p < .025$], but more importantly, both the effect of groups and the interaction were clearly nonsignificant (Fs < 1). Any differential depression of matching scores among the three groups should not then be attributable to differences in error occurrences in recall. Pearson product-moment correlations between incorrect responses in modified free recall and correct responses in associative matching were ($df = 14$) $-.53$ (Group IM, $p < .025$), $+.07$ (Group NIT, chance), and $-.17$ (Group ITT, chance). The matching performance of Group IM thus was negatively related to errors during recall, a result consistent with Keppel's findings. However, the inconsequential correlations in Groups ITT and NIT indicate that the scores of these subjects were relatively unaffected by their performance on recall. The present estimate of retroactive inhibition of specific associations in Group ITT relative to Group IM therefore appears a conservative one. It is conceivable that the depression effect observed by Keppel and in the IM group of the present study may not occur when either relevant or irrelevant interpolated activity intervenes between acquisition and recall.

---

[11] G. Keppel, Association by contiguity: Role of response availability, *J. exp. Psychol.*, 71, 1966, 624–628.

The explanation of the loss of specific associations in both the paired-associate and verbal-discrimination tasks poses a challenging theoretical problem since (*a*) the specific associations of two unrelated lists presumably do not enter into an A–B, A–C relationship and (*b*) the effects of the unlearning of context-item associations are assumed to be overcome through the use of the associative-matching technique. The present study was not designed to test the several theoretical explanations that have been suggested.[12] The present data do, however, appear relevant to Garskof's suggestion that the nominal C–D list supplied by the experimenter may approach a functional A–C list for the subject and thus account for A–B unlearning. Such an explanation may be tenable with the nonsense-syllable stimuli used by Garskof and by Houston, since there typically exists with nonsense-syllable materials an overlap in letter content. The nominal stimulus-letter overlap and idiosyncratic similarities discovered by the subjects could lead to a functional A–C list. The present study, however, used preexperimentally integrated, semantically dissimilar adjectives as stimuli. Under these conditions it seems unlikely that stimulus selection would occur in such a manner as to produce a functional A–C relationship.

The authors of the present study favor an explanation based on the competition at recall between the various kinds of relevant contextual associations formed during List 1 and List 2 learning. That is, context–specific-association bonds (context–W→R and context–R→W) of the two lists may enter into competition not only during List 2 learning but also during performance on the associative-matching tasks, to the extent that the contextual cues during recall continue to elicit the (now higher in strength) W_2–R_2 associations. Future research on the specific-association–unlearning phenomena might beneficially include signal-detection approaches as a more sensitive technique for assessing the subtle processes that are apparently involved.

---

[12] Garskof (see n. 3 above).

SOME EFFECTS OF STIMULUS-EXPOSURE TIME ON CHOICE-REACTION TIME

RICHARD E. CHRIST, *Kansas State University*

Abstract. Eighty-six college students were run in two experiments investigating choice-reaction time as a function of mixed and constant exposure time. Within the range of exposure times explored (50–150 msec.), conditions employing a constant exposure time did not differ from one another. However, when a series of short exposure times was mixed with a series of intermediate exposure times, the reaction time to both was increased. The major conclusion is that stimulus and subject factors develop within the experimental situation to produce effects that cannot be predicted from knowledge of only the stimulus-exposure conditions.

There have been a number of recent studies on the effects of stimulus-exposure time on reaction time. Raab and his associates have shown that simple reaction time became shorter with increasing exposure time, but only for brief exposures (up to about 20 msec.) and only at dim luminances (up to about 30 mL.)[1] With more complex discrimination tasks, Kaswan and Young have shown that both accuracy and choice-reaction time were affected by a much wider range of exposure times (4–512 msec.). Accuracy increased with increasing exposure time at all luminance levels but was largely unaffected by variations in luminance (.09–11.84 mL.). Choice-reaction time varied as a function of both luminance and exposure time; the relation of luminance and exposure time to choice-reaction time gradually shifted from direct at dim and brief exposures, to inverse at bright and long exposures.[2]

Various procedural problems can arise, however, when at-

Received for publication January 5, 1970. This research was supported by Faculty Research Grant F-362 from the Bureau of General Research, Kansas State University. Special thanks must be given to David Strohm for assistance in the collection of data. Mr. Strohm was supported by the NSF Undergraduate Research Participation Program, Grant GY-13.

[1] D. Raab, E. Fehrer, and M. Hershenson, Visual reaction time and the Broca-Sulzer Phenomenon, *J. exp. Psychol.*, 61, 1961, 193–199; D. Raab and E. Fehrer, The effect of stimulus duration and luminance on visual reaction time, *J. exp. Psychol.*, 64, 1962, 326–327.

[2] J. Kaswan and S. Young, Stimulus exposure time, brightness, and spatial factors as determinants of visual perception, *J. exp. Psychol.*, 65, 1963, 113–123; J. Kaswan and S. Young, Effect of luminance, exposure duration, and task complexity on reaction time, *J. exp. Psychol.*, 69, 1965, 393–400.

tempts are made to relate the speed of reaction to stimulus variables. Raab et al. have shown that the effect found for stimulus duration in simple reaction tasks was in part a function of the method employed for presenting various levels of the exposure-time variable.[3] They concluded, as did both Johnson and Jarl,[4] that when a short-duration flash is presented repeatedly over a block of trials, various attitudinal factors or preparatory adjustments may develop that arouse the subject to react as fast as possible, while a block of longer durations might encourage some relaxation of effort. To assess the influence of exposure time without possible confounding by these differential sets or degrees of effort, Raab et al. concluded, the different levels of exposure time should be presented in a random sequence.

The use of different exposure times within a block of trials may, however, introduce yet another procedural problem. Christ and Gerjuoy reported that the relative difficulty of additional choice responses, made in an otherwise common experimental situation, affected the reaction time to the common stimuli.[5] Specifically, these experimenters required their subjects to press one of nine buttons in response to arabic numerals. The stimuli were presented to one group of subjects for exposure times of 50, 60, 80, 100, and 120 msec. For another group of subjects, a 140-msec. duration was substituted for the 50-msec. one in the same random sequence of stimuli and exposure times. The results showed that after some practice, the choice-reaction time to the common set of stimulus-exposure times was shorter when the additional choice responses were made to stimuli exposed for relatively long times (140 msec.) than for shorter times (50 msec.). The experimenters concluded that the 140-msec. exposure times created a context in which the entire task seemed easier.

The purpose of the following experiments was to further investigate the influence of exposure time on a context-difficulty factor in a choice-reaction task. Experiment I was designed to replicate the Christ and Gerjuoy results, but in addition, task difficulty was manipulated by varying the number of alternative stimuli.

---

[3] Raab, Fehrer, and Hershenson (see n. 1 above).

[4] H. M. Johnson, Reaction time measurements, *Psychol. Bull.*, 20, 1923, 562–589; V. C. Jarl, Method of stimulus presentation as antecedent variable in reaction time experiments, *Acta Psychol.*, 13, 1957, 225–241.

[5] R. E. Christ and H. Gerjuoy, Difficulty context and practice effects in disjunctive reactions, paper presented at the meeting of the Midwestern Psychological Association, Chicago, 1965.

EXPERIMENT I

Method

Apparatus. A Gerbrands two-field tachistoscope was used to present Deca-Dry rub-on numerals (1, 2, 3, and 4). These stimuli were 1.3 cm. high with a .2-cm. stroke width. Each stimulus was individually presented in the center of a 7.6 cm. × 2.2 cm. white field, which was in turn centered in a 18.4-cm. square black field. The interstimulus field of the tachistoscope was identical to the stimulus field, but the white field was blank. The white rectangular ground in both fields had a luminance of 17.56 mL.

A Mallory Sonalert (4500 Hz.) sounded 1 sec. before and terminated .5 sec. prior to the presentation of each stimulus and served as a ready signal. Hunter timers were used to control the duration of the ready signal, the lag between the ready signal and the stimulus, and the lag (8 sec.) between *S*'s response and the beginning of the next trial. One of three additional timers, selected by means of a three-position switch, was used to control the duration of the stimulus exposure.

The *S* responded by pressing one of two or one of four telegraph keys; two placed near his right hand and two near his left hand. These keys were wired to four response-indicator lights and to a Standard Electric clock which measured choice-response time to within .01 sec.

Procedure. Thirty-six students volunteered to participate individually for 240 trials. Sixteen *S*s served in a two-choice reaction task; they were instructed to press the extreme left key when they saw the numeral 1 and the extreme right key when they saw the numeral 2. The remaining 20 *S*s served in a four-choice reaction task in which the extreme left and center left keys corresponded to the numerals 1 and 2 respectively, and the center right and extreme right keys corresponded to the numerals 3 and 4. Half of the *S*s in each choice condition (Group S) had the stimuli presented for 50 or 100 msec; for the others (Group L), 150 msec. was substituted for 50 msec. in the same stimulus sequence. The order of stimuli and the durations were randomized with the restriction that each stimulus and each exposure time occur equally often in each of five 48-trial blocks. All *S*s were instructed to respond as rapidly and as accurately as possible. Each *S* was given eight practice trials using both the common and additional exposure times to ensure that he understood the task.

Results

The median choice-reaction time for each duration within each trial block was determined for each subject. The top portion of Table I shows the means of these individual medians averaged over blocks and subjects. Table I also shows, in parentheses, the standard deviations of the average choice-reaction times within the groups. Errors were extremely infrequent for all subjects, and groups did not differ in terms of the frequency of errors.

In the two-choice condition there was a slight tendency over trial blocks for a shorter reaction time to the 100-msec. stimulus if the additional exposure time was relatively long (Group L) than if it was relatively short (Group S). However, analyses of the data showed that the difference between groups was not significant [F (1, 14) < 1.0]. Likewise, there were no significant effects attributable to blocks [F (4, 56) $= 2.42, p > .05$] or to the interaction of groups and blocks [F (4, 56) < 1.0]. Similarly, no significant effects were found for exposure time within Group S [F (1, 7) < 1.0] or Group L [F (1, 7) $= 2.37, p > .05$].

TABLE I

MEANS AND STANDARD DEVIATIONS OF CHOICE-REACTION TIMES FOR EACH STIMULUS-EXPOSURE TIME WITHIN TREATMENT GROUPS

	Exposure time		
Treatment groups	50	100	150
Experiment I			
Two-choice			
Group S	388 (29)	386 (38)	
Group L		376 (70)	381 (74)
Four-choice			
Group S	538 (47)	542 (46)	
Group L		499 (43)	513 (45)
Experiment II			
Four-choice			
Group S	547 (56)	538 (56)	
Group L		494 (47)	498 (48)
Group K50	496 (38)		
Group K100		494 (51)	
Group K150			499 (42)

Note: All times in msec. Standard deviations shown within parentheses.

In the four-choice conditions, subjects in Group L consistently responded more rapidly to the 100-msec. exposure time than did subjects in Group S. A one-tailed uncorrelated t test was used to test the difference between groups using each subject's mean median choice-reaction time as data. This comparison showed the overall difference between groups to be significant [t (18) $= 2.03$, $p < .05$]. An analysis of variance performed on the data over blocks showed a significant practice effect [F (4, 72) $= 3.04, p < .05$], but the choice-reaction times of both groups decreased over blocks at the same rate, as is suggested by a nonsignificant interaction of groups and blocks [F (4, 72) < 1.0]. No significant effects were found for exposure time within Group S [F (1, 9) < 1.0] or Group L [F (1, 9) $= 4.93, p > .05$].

EXPERIMENT II

Since the effects of mixed exposure time found in Experiment I were relatively small, it was decided that the replicability of these results needed to be established. This was one purpose of Experiment II. In addition, the data from Experiment I do not allow one to determine if the mixed exposure times caused an increase in the choice-reaction time of Group S, a decrease in the choice-reaction time of Group I, or both, relative to that expected without the mixing of exposure times. Hence, Experiment II also employed separate groups run at each level of exposure time as a control for the context effects created by combining exposure times within an experimental session. Finally, since the context effects were present throughout the five 48-trial blocks in Experiment I, smaller blocks of trials were employed in Experiment II to assess the development of the context effect.

Method

Apparatus. The apparatus was that used in Experiment I.

Procedure. Fifty students participated individually for 96 four-choice reaction trials. Ten *Ss* each served in Groups S and L, defined as in Experiment I. The remaining 30 *Ss* were assigned, 10 each, to one of three constant exposure-time conditions: Group K50 had 50-msec. exposure times, Group K100 had 100-msec. exposure times, and Group K150 had 150-msec. exposure times. The same random sequence of stimuli was used for all groups and each stimulus, and each combination of stimulus and exposure time for Groups S and L occurred equally often within each of six 16-trial blocks. Instructions and practice were as in Experiment I.

Results

Again the median choice-reaction time was derived for each combination of exposure time and trial block for each subject in Groups S and L. For the three constant exposure-time groups, the median choice-reaction time was derived separately for the trials within blocks that corresponded to the common and additional exposure-time trials for Groups S and L. However, since no within-group differences were found for these latter data, only the data from comparable trials were used in analyses between the mixed and constant exposure-time groups. The lower portion of Table I shows the means of these individual medians averaged over blocks and subjects within groups, and gives the standard deviations of the average choice-reaction times. Errors were again very infrequent, and groups did not differ in terms of the frequency of errors.

As in Experiment I, there was a consistent tendency over trials for subjects in Group L to respond more rapidly to the common exposure time than did subjects in Group S. A one-tailed comparison between groups, using each subject's average median as data, was significant [$t\ (18) = 1.80, p < .05$]. Again, however, an analysis of variance failed to show an overall effect of blocks [$F\ (5, 90) = 1.03, p > .05$] or an interaction of groups and blocks [$F\ (5, 90) < 1.0$]. As in Experiment I, no significant differences were found between durations within Group S [$F\ (1, 9) = 3.10, p > .05$] or Group L [$F\ (1, 9) < 1.0$].

For the constant exposure-time groups there were no between-group and no trial-block effects. Separate comparisons between comparable exposure-time conditions in the mixed and constant groups showed that Group 50K and Group 100K responded faster than Group S at comparable levels of exposure time [$t\ (18) < 2.53$ and 1.73 respectively, $p < .05$ in each case]. Comparisons between the constant exposure-time groups and Group L were not significant [$t\ (18) < 1.0$].

DISCUSSION

The main finding of the present experiments was the consistency in results over independent investigations for the mixed exposure-time groups. If it is assumed that exposure time over the range explored has no independent effect on choice-reaction time—and the data from the constant exposure-time condition in Experiment II suggest this is the case—then the probability of obtaining by chance four means such that the two from Group S are longer than the two from Group L is .167. The probability of obtaining this same organization of mean choice-reaction times in each of the three independent studies is less than .005. Hence, although the differences in choice-reaction time favoring Group L were very small, the replicability of these findings was highly significant, and it must be concluded that some sort of context effect in the mixed exposure-time conditions was responsible for the obtained results.

The results obtained for the four-choice, mixed exposure-time groups substantiate the earlier findings of Christ and Gerjuoy (see n. 5). Specifically, subjects responded more quickly to stimuli presented at a moderately long exposure time if other, additional stimuli were exposed for relatively long times rather than for relatively short times. Unlike the Christ and Gerjuoy findings, the difference between groups did not develop gradually over blocks of trials but rather was present throughout the experiment. This discrepancy between the two studies may be due to the practice given subjects in this study before the start of the experiment. No such prior practice was given subjects in the Christ and Gerjuoy study.

Another reason for the different effects found for trial blocks in the two studies may be the number of alternative stimuli employed. Indeed, the results found in this study for numbers of alternative stimuli suggest that the number of alternative stimuli may be a major factor in establishing differential responding to the common exposure time. This in turn would suggest that a context-difficulty factor, based on a random sequence of exposure-time conditions, cannot be established unless the task is already relatively difficult as a result of the number of alternative stimuli.

If the choice-reaction times of the constant groups do in fact reflect the independent effects of exposure time within the range tested, the shift in choice-reaction time for Group S could then reflect a change in the apparent ease of the task—a change which developed as a result of the contrast between the moderately long and short exposure times. If, however, as Kaswan and Young reported,[6] the relative ease of visual discriminations is directly related to exposure time, the results of the mixed exposure-time groups in both Experiments I and II could be accounted for by an assimilation process wherein subjects in Group S would encounter a more difficult task than would subjects in Group L. In this latter case, the consistently fast choice-reaction times obtained for the constant exposure-time groups in Experiment II could be attributed to differential sets which develop from the constant stimulating conditions.[7]

---

[6] Kaswan and Young (see the first reference in n. 2 above).

[7] See Jarl (n. 4 above); Johnson (n. 4); and Raab, Fehrer, and Hershenson (n. 1 above).

Either of the above explanations for the obtained results could be supported from previous work in this area. The present two experiments cannot, however, distinguish between these two interpretations. Furthermore, the present experiments cannot determine whether the observed changes in choice-reaction time were due to a change in the ease of discriminating stimuli or to a change in the readiness to respond. The major conclusion is, rather, that measurements of reaction time are very sensitive to variations in procedure. Before any effects observed in these measurements can be attributed to stimulus variables or to subject variables, much more must be known about the influence of factors that develop within the experimental situation.

FALSE RECOGNITIONS AS A FUNCTION OF NUMBER OF PRESENTATIONS

JAMES W. HALL and EDWARD E. KOZLOFF, *Northwestern University*

Abstract. A list of words was presented visually to 160 college students, who were asked to indicate whether each word had or had not appeared earlier on the list. Late in the list were associates of those 'critical stimulus' words that had appeared earlier, either one, three, five, or seven times. The frequency of false recognition of the associates first increased, then declined, as a function of the number of times their corresponding critical stimulus words had occurred, and was greater than that for nonassociates of previously appearing words. These results are considered in terms of a two-stage recognition model wherein the perceived situational frequency of a word is critical in the recognition decision.

Several word-recognition studies have shown that the frequency of false recognition of words associatively related to previously appearing words is higher than for words not so related.[1] This phenomenon has been interpreted as providing evidence for associative encoding for memory.[2] That is, when a word is presented for learning, words associatively related to the word to be learned are elicited as implicit responses. These *implicit associative responses* (IARs) are encoded and stored along with the word to be learned, frequently leading to subsequent false recognitions.

In earlier papers it was suggested that the perceived situational frequency of words presented for recognition is a critical cue for recognition decisions.[3] This assumes, as proposed by Ekstrand, Wallace, and Underwood, that the frequency with which a word occurs is recorded in some manner.[4] The perceived situational frequency of a word presented for recognition is a function of the number of prior presentations of the word (perceptions of the

---

Received for publication November 4, 1969. The study was supported by a grant to the senior author from the United States Office of Education.

[1] For example, see G. A. Kimble, Mediating associations, *J. exp. Psychol.*, 76, 1968, 263–266.

[2] B. J. Underwood, Attributes of memory, *Psychol. Rev.*, 1969, in press.

[3] J. W. Hall, Errors in word recognition and discrimination by children of two age levels, *J. educ. Psychol.*, 60, 1969, 144–147; J. Hall, R. Sekuler, and W. Cushman, Effects of IAR occurrence on response time during subsequent recognition, *J. exp. Psychol.*, 79, 1969, 39–42.

[4] B. R. Ekstrand, W. P. Wallace, and B. J. Underwood, A frequency theory of verbal discrimination learning, *Psychol. Rev.*, 73, 1966, 566–578.

word by the subject) plus the number of implicit occurrences of the word, as a result, for example, of rehearsal of the words. Thus, situational-frequency counts above zero exist either when a word was previously presented or when it occurred solely as an IAR to a presented word. Presumably the count would be higher for an actually presented word, since such a word normally is perceived by a subject, pronounced implicitly, and perhaps rehearsed. One would expect less attention to be given to any particular IAR.

Hall, Sekuler, and Cushman proposed a model to account for performance when words presented for recognition fall into one of three classes: words presented earlier ('old' words), 'new' words that are strong associates of previously presented words, and 'new' words unrelated to presented words.[5] According to this model, the subject establishes a perceived-situational-frequency criterion, against which he judges words when recognition decisions are required. That is, words whose perceived situational frequencies exceed that criterion are judged 'old' and those whose do not are judged 'new.' Also, the further above the criterion, the greater the confidence and the speed with which they are judged 'old.'

The above analysis leads to a clear-cut prediction about the relationship between the number of occurrences of a word and the probability that the word will be judged 'old' when it is presented for recognition. First, the greater the number of presentations of a word, other things equal, the greater the probability of that word subsequently being correctly judged 'old.' Second, the greater the number of presentations of a word that consistently elicits a particular associate as an IAR, the greater the probability of that associate being incorrectly judged 'old.' The validity of the first of these two hypotheses needs no additional documentation. It is the second with which the present experiment is concerned. Some evidence relevant to it comes from an experiment by Underwood in which a set of words was presented either one or three times prior to presentation of strong associates of these words.[6] False recognitions of the associates were more frequent for the three-presentation than for the single-presentation condition. Such a difference seems reasonable in that one would not expect any word to elicit the same IAR for all subjects at its

[5] Hall, Sekuler, and Cushman (see n. 3 above).

[6] B. J. Underwood, False recognition produced by implicit verbal responses, *J. exp. Psychol.*, 70, 1965, 122–129.

first presentation. However, a second and third presentation might well result in nearly 100% occurrence of strong associates as IARs, thus providing increasing basis for subsequent false recognitions. However, the prediction following from the Hall, Sekuler, and Cushman analysis (see n. 3) is that with increasing numbers of occurrences as IARs of a set of words, the perceived situational frequencies of those words eventually would be shifted upward to the point at which all would be falsely recognized, a possibility that appeared intuitively unlikely. So the prediction was tested in the experiment reported here by presenting a set of words either one, three, five, or seven times (within-subjects) before presenting the associates of these words. Recognition performance for these conditions and for repeated and control words was then examined.

METHOD

Subjects. Ss were 160 students enrolled in an introductory psychology course at Northwestern University.

Design and procedure. A list of 99 words was presented to the *Ss*. The first 66 items in the list were 7 filler (F) words, 9 repeated (R) words, and 12 critical stimulus (CS) words. The F words were inserted to increase both the difficulty and face validity of the task. The R words were items which were repeated among the final 33 words in the list. Each of the 12 CS words was assumed, on the basis of normative word-association data, to elicit a particular response as an IAR and to do so with relatively high frequency. The CS words were divided into four groups of three words each. One group occurred once (CS1), one group occurred three times (CS3), one group occurred five times (CS5), and the remaining group seven times (CS7). Thus, there were four conditions based on the number of presentations of the CS words. A counterbalancing procedure was employed so that each group of three CS words occurred under each condition, resulting in four different forms of the list with one-fourth of the *Ss* randomly assigned to each form.

The remaining 33 words were exactly the same for all lists. In addition to the 9 repeated (R) words were 12 experimental (E) words and 12 control (C) words. The 12 E words were strong natural-language associates of the 12 CS words that had occurred previously. These words are designated E1, E3, E5, or E7, depending on the number of prior presentations of the CS words presumed to elicit them. The CS-E pairs were selected on the basis of published word-association data.[7] The C words were approximately equivalent to the E words in general frequency of occurrence according to the Thorndike and Lorge norms,[8] but they were not strong associates of any other words in the list. Table I shows one form of the list and the function of each word.

---

[7] D. S. Palermo and J. J. Jenkins, *Word Association Norms: Grade School through College,* 1964.

TABLE I

Word List and Function of Each Item

king	CS	hard	CS	train	R
on	CS	brush	R	fast	E
lamp	CS	slow	CS	bath	C
hand	R	butter	CS	brush	R
on	CS	king	CS	flower	E
mountain	F	city	R	shoe	C
king	CS	news	R	wild	R
lamp	CS	slow	CS	quiet	R
month	F	hard	CS	soft	E
blossom	CS	lamp	CS	rough	C
on	CS	hammer	CS	bread	E
lamp	CS	butter	CS	sleep	E
slow	CS	quiet	R	old	E
hammer	CS	on	CS	whiskey	C
king	CS	blossom	CS	pretty	C
door	F	king	CS	cut	E
friend	R	table	CS	between	C
hand	R	music	R	wing	C
train	R	scissors	CS	hand	R
lamp	CS	table	CS	light	E
slow	CS	blossom	CS	news	R
blossom	CS	on	CS	hat	C
wild	R	king	CS	name	C
on	CS	slow	CS	cow	C
king	CS	butter	CS	city	R
hammer	CS	young	CS	clear	C
clock	F	spider	F	chair	E
friend	R	bed	CS	nail	E
table	CS	hard	CS	music	R
hammer	CS	hammer	CS	queen	E
on	CS	lamp	CS	actor	C
lamp	CS	sad	F	friend	R
blossom	CS	apple	F	off	E

The lists were presented visually by means of a slide projector to groups of from four to eight *Ss*. The presentation rate was 2.5 sec. per word. The *Ss* were instructed to indicate 'yes' on an answer sheet if a word had occurred earlier in the list and 'no' if it had not. Only responses to the final 33 were relevant to the purposes of the experiment, so only those items were scored and included in the results.

---

[8] E. L. Thorndike and I. Lorge, *The Teacher's Word Book of 30,000 Words*, 1944.

RESULTS

The mean percentages of 'old' responses were 82.8 for the repeated words, 7.5 for the experimental words (combining the four conditions E1, E3, E5, and E7), and 2.7 for the control words. The mean percentages of recognitions that were false recognitions for the various categories of experimental words and for the control words are shown in Figure 1. The frequency of false recognitions increased sharply (from one to three presentations of the critical stimulus words), then fell off with additional presentations; and the difference among the four experimental-word means was reliable [$F\ (3, 636) = 3.09,\ p < .05$]. The overall differences between experimental and control words in frequency of false recognitions was highly reliable [$t\ (159) = 6.88,\ p < .01$]. Even in the case of E7 words the proportion of false recognitions was reliably greater than for the control words [$t\ (159) = 3.98,\ p < .01$]. For this last-mentioned test, the difference in mean proportion of false recognitions was obtained by multiplying the mean number of false recognitions of E7 words by four, then subtracting the mean number of false recognitions of the control words. The error term employed was that used in testing the overall difference between experimental and control words.

DISCUSSION

In an earlier paper it was suggested that subjects make their recognition decisions on the basis of the perceived frequency of a word in relation to a perceived-frequency criterion set by the subjects.[9] On the average, the perceived frequency of repeated words would be higher than that for experimental and control words because every repeated word had been presented to every subject earlier. The distribution of experimental words in terms of perceived frequency would be higher than for control words, because of the prior occurrence of experimental words as IARs, but not as high as for repeated words, since it is assumed that not all experimental words do occur as IARs for all subjects. This analysis is consistent with the differences in frequency of 'old' responses for the control, E1, E3, and repeated words in the present experiment. In particular it predicts the observed higher frequency of false

---

[9] Hall, Sekuler, and Cushman (see n. 3).

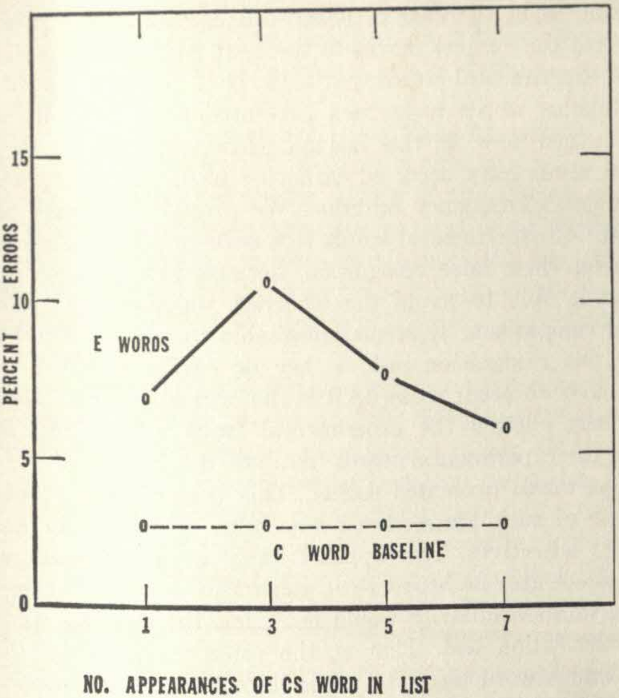


FIG. 1. Percent false recognitions of experimental (E) and control (C) words as a function of the number of appearances of the critical stimulus (CS) word earlier in the list

recognitions of E3 than of E1 words, assuming that the two additional occurrences of the CS3 words resulted in the occurrence as IARs of an increasing number of the E3 words, these then being shifted upward in perceived frequency past the perceived-frequency criterion. However, according to this analysis, the perceived frequency of E5 and E7 words, and thus the frequency of 'old' responses, would be even greater, rather than less as was found. In other words, this recognition-performance model simply fails to account for a substantial portion of the results in this experiment. Rather than discarding the model completely at this point, the authors are led to the following modification of it.

A two-stage recognition model is proposed, the first stage consisting of a comparison of the perceived situational frequency of the word with the subject's perceived-situational-frequency criterion, as in the earlier analysis. When a word falls below that criterion, as in the case of nearly all control words, it is judged 'new' and the subject moves to the next item. A substantial number of experimental words, particularly if the corresponding critical stimulus words have been presented only once, are immediately judged 'new' in this fashion. However, many of the experimental words have occurred earlier as IARs and so register above the subject's frequency criterion. We propose that such prior occurrence of experimental words is a necessary but insufficient condition for their false recognition, because in most such cases the subject is able to avoid the incorrect judgment 'old' through a second comparison. It seems reasonable to assume that IARs occur during recognition just as they do earlier and that the word most likely to occur as an IAR is the critical stimulus word which previously elicited the experimental word as an IAR. In other terms, the experimental words 'remind' the subject of the critical stimulus words presented earlier. This is particularly probable in the case of such bidirectional associates as *table-chair* and pairs of polar adjectives. This appears likely even with such unidirectional associates as *scissors-cut* and *blossom-flower*, where the response words ordinarily would not elicit the stimulus words on a word-association test. That is, the prior occurrence of the critical stimulus word ought to prime its subsequent occurrence as an IAR. The priming effect would be particularly powerful when the critical stimulus word had occurred a number of times previously, as did the CS5 and CS7 words in the present experiment. Moreover, if the experimental word had been elicited as an IAR by the critical stimulus word earlier, then through backward association the probability of the critical stimulus word occurring as an IAR when the experimental word was presented for recognition would be increased.

Once the subject is reminded in the above fashion of the critical stimulus word, the decision as to which of the two words really did occur earlier can be analyzed as a verbal-discrimination problem, and the frequency theory proposed by Ekstrand, Wallace, and Underwood (see n. 4) can be applied. This theory states that the major cue for discrimination is a disparity in perceived frequency between the two items. In experiments such as the present one, a difference in perceived situational frequency between critical stimulus words and whatever words occur as IARs is presumed due to the subject's relatively greater attention to and implicit rehearsal of any word actually presented. This disparity would increase as the number of presentations of a given critical stimulus word increases. It is primarily this disparity in perceived situational frequency between a critical stimulus word and its corresponding experimental word that would allow the subject to generally select the former and avoid false recognition of the latter. The fact that IAR false recognitions do occur in some instances may be due either to the failure of the critical stimulus word to come to mind when the experimental word is presented or to insufficient difference in perceived situational frequency between the two. The probabilities of both of these events would be decreased as the number of occurrences of the critical stimulus word is increased.

The notion that subjects discriminate between experimental and critical stimulus words on the basis of the difference between them in perceived situational frequency leads to a rather clear-cut prediction. Suppose that through word-association testing an experimenter identifies three words not strongly associated with one another but each of which elicits the same word as a strong associate. For example, *lamp, heavy,* and *dark* each elicit *light* with high frequency but are very weakly associated with one another. Suppose further that under one condition subjects are presented only one of these three words but with that word occurring three times in the list, whereas the remaining subjects are presented each of the three words a single time. In the first instance, each time a 'frequency count' is added to the experimental word (*light,* in this case), an even greater frequency count is added to the critical stimulus word (say, *lamp*). Thus, a basis for discrimination between the experimental and critical stimulus words in terms of the discrepancy in perceived situational frequency is developed. In the second instance, where converging associations are involved, there is no build-up in the frequency difference with additional occurrences of the experimental word because it is a different critical stimulus word that occurs each time. Thus, one would predict more difficult discrimination and a correspondingly higher frequency of false recognition in the case of converging associates.

NOTES AND DISCUSSION

A LINE ILLUSION WITH IRRELEVANT DEPTH CUES

Many attempts have been made to refute Gregory's hypothesis that the Müller-Lyer illusion, among others, arises when depth cues are present as a consequence of inappropriate constancy scaling.[1] Thus, for example, Zanforlin has presented many instances of figures that give rise to impressions at deviance with plausible predictions from the theory.[2] Since the theory appeals both to 'primary,' or depth-cue, scaling and to 'secondary,' or hypothesis, scaling,[3] interpretation becomes too complex to admit of disproof by isolated instances; indeed, it has been argued by Humphrey and Morgan that the theory is essentially untestable.[4]

However, the accumulation of contrary instances, where illusory effects are obtained in circumstances where depth cues are clearly irrelevant, must eventually lead either to the abandonment of the theory or at least to a modification of the role assigned to depth cues. Attention has recently been drawn to one such negative instance, a derivative of the Müller-Lyer figure in which a line is bounded by one ingoing and one outgoing arrowhead.[5] In this figure, a mark located at the midpoint of the line appears to divide the line into two *unequal* portions. The apparent inequality persists regardless of which of the two possible three-dimensional interpretations a subject gives to the figure. Depth cues thus appear irrelevant to the illusion, which nevertheless contains the essential features of the Müller-Lyer.

The intention of the present note is to demonstrate a related modification of the Müller-Lyer, in which depth cues appear irrelevant, and in which the nature of the illusory effect is different. The new effect is shown in Figure 1. Observe that while there is, of course, no difference in the apparent length of the two lines, there is an apparent shift in position.

---

[1] R. L. Gregory, Distortion of visual space as inappropriate constancy scaling, *Nature*, 199, 1963, 679–680.

[2] M. Zanforlin, Some observations on Gregory's theory of perceptual illusions, *Quart. J. exp. Psychol.*, 19, 1967, 193–197.

[3] R. L. Gregory, Comments on the inappropriate constancy scaling theory of the illusions and its implications, *Quart. J. exp. Psychol.*, 19, 1967, 219–223.

[4] N. K. Humphrey and M. J. Morgan, Constancy and the geometric illusions, *Nature*, 206, 1965, 744–745.

[5] M. J. Morgan, Estimates of length in a modified Müller-Lyer figure, this JOURNAL, 82, 1969, 380–384.

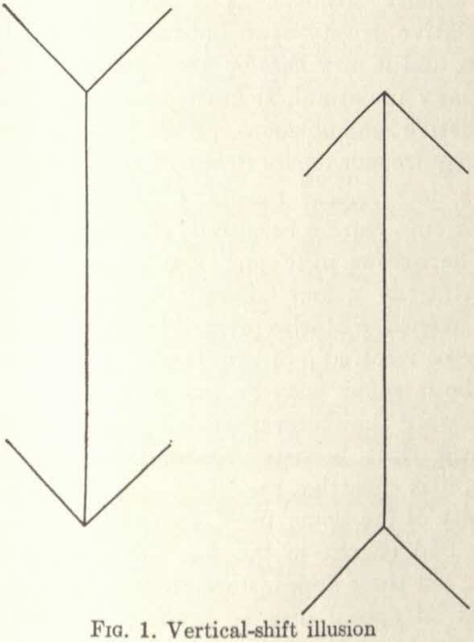


FIG. 1. Vertical-shift illusion

Clearly, the apparent downward shift in location of the right-hand arrow, and upward displacement of the left, may be used to reconstruct the Müller-Lyer effect. Thus, if the downward-displaced arrowhead at lower right were transferred to the left, while the upward-displaced arrowhead at lower left were transferred to the right, a conventional Müller-Lyer figure would be restored; the restored figure should then show a discrepancy in apparent line length equal to twice the amount of the apparent shift. The Müller-Lyer effect may therefore be due to the same perceptual mechanism which gives rise to the vertical shift.

The apparent shift itself is most probably explained by a process of visual averaging. That is to say, judgments of length are normally made on the basis of the total complex of end cues, the end point being taken as a form of weighted mean of the line and arrowhead elements. A related explanation for the Müller-Lyer, characterized as the 'limited acuity' theory, was discarded by Gregory on the grounds that it is perceptually quite possible to separate line from arrowheads.[6] However, it is not necessary to assert that subjects cannot perform the separation, merely that they normally do not. Asking humans to make pure judgments of any physical dimension is usually artificial, as is shown dramatically by the influence of relative density upon judgments of weight in the size-weight illusion, and it may here be assumed that judgments of line length are equally unnatural. It must, of course, be admitted that in order to derive unambiguous predictions, a visual-averaging theory would require much more stringent specification.

However, it does appear that any alternative explanation in terms of depth cues can be ruled out, since these cues are doubly ambiguous. The arrows of Figure 1 may be given three-dimensional interpretations in four different ways. Each element of the figure may be likened to Mach's reversible open-book illusion.[7] Thus, the figure may be resolved into two 'books' with the 'spines' toward the observer, both 'spines' may be seen as receding, or either 'book' may project toward the observer while the other recedes. Since the illusory vertical shift persists regardless of the applied depth interpretation, it is clear that the effect is not mediated by the perceived proximity of the 'spine' lines. Two of the possible interpretations produce odd effects, in the form of counterrotation of the apparent axes, but these appear unrelated to the shift phenomenon. It is true that an appeal may be made to a form of hypothesis scaling, since the 'pages' may be interpreted as lying at different locations along the same plane at right angles to the line of sight of the observer; but this explanation seems unnecessarily elaborate and is, in any case, different from an explanation in terms of differential apparent depths of the 'spines.'

The illusion figure shown seems therefore to present a further difficulty for a theory based on depth perception and to favor a simpler explanation—perhaps one based on an averaging process. Since the present figure may be transposed into the Müller-Lyer effect, it seems probable that the Müller-Lyer itself would be better served by something other than a depth theory, using presented rather than inferred features.

DENNIS H. HOLDING, *University of Louisville*

---

[6] R. L. Gregory, *Eye and Brain*, 1966, 141–142.
[7] M. Luckiesh, *Visual Illusions*, (1922) 1965, 69.

A VARIABLE-INTERVAL GENERATOR WITH AUTOMATED CHANGE IN MEAN INTERVAL

The most commonly employed variable-interval generator consists of a microswitch arm riding on the surface of a moving loop of 16-mm. film in which irregularly spaced holes are punched. The arm drops into these holes, providing a brief contact closure. The film is driven at a constant speed by an a.c. motor, which results in a series of pulses with the mean interval between pulses determined by the spacing of holes.

It often is desirable to change the mean interval between pulses in the course of an experiment. Where the change need only be made between sessions, one loop of film may be replaced by another in which the holes are spaced differently. Where the length of the mean interval must be changed during an experimental session, the usual practice is to shift from one generator to another, a solution which requires as many generators as mean intervals.

A more flexible and much more economical procedure is to use a generator driven by a d.c. motor, whose speed of rotation is, of course, directly proportional to the voltage applied. Changes in speed and therefore in the mean interval then can be programmed within a session by routing the current operating the motor through resistors of different value. The mean value of motor-driven variable-ratio programmers may be altered within sessions by the same means.

JOHN BINTZ, *University of California, Davis*

The development of this device was supported by Grant NH-17032 from the Public Health Service.

AN UNOBTRUSIVE PHOTOCONDUCTIVE EYEBLINK TRANSDUCER FOR HUMAN SUBJECTS

Eyeblinks can be unobtrusively recorded in ambient illumination (overhead fluorescent and incandescent) with a specially mounted photocell which, properly positioned and connected, registers differences in light reflected from an open or closed eye.

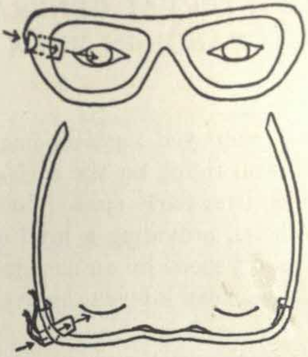


FIG. 1. Front and top views, showing placement of the photocell sleeve in the frame

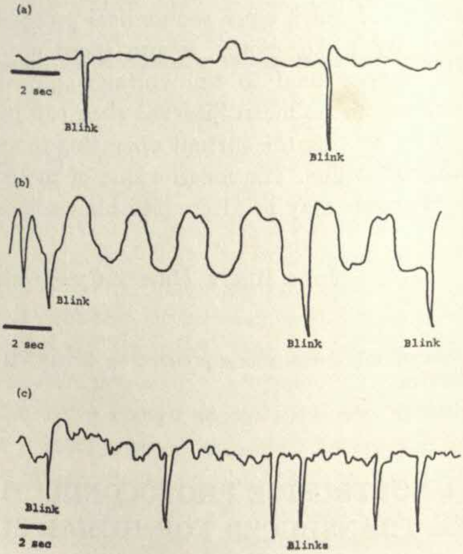


FIG. 2. A record of eyeblinks from a subject (a) staring, (b) performing instructed lateral movements, and (c) reading

A Type 4 CdSe cell such as the Clairex CL6040 or its equivalent is wired as one arm of a Wheatstone bridge. After balancing, the output of the bridge is fed directly to a pen recorder for amplification. The Beckman 9801 d.c. coupler (or 9806A) and RS Dynograph provide suitable records. A pair of sunglasses with the lenses removed holds the cell, as shown in Figure 1. The cell is recessed in a blackened aluminum sleeve, which serves to eliminate extraneous light. The placement of the sleeve unit in the frame is critical, of course, if there is to be no visual obstruction for the wearer. A record of eyeblinks from the transducer is shown in Figure 2. Blinks are readily distinguishable from such artifacts as head and eye movement.

The transducer may also prove useful for those working with eyelid conditioning. By the addition of another sleeve, placed concentrically about the photocell sleeve, an air puff could be delivered to the eye. The whole unit would be easily put on and light in weight.

HENRY MICHALEWSKI, *California State College at Long Beach*

CORRECTION

The listing for the authors of "Agreement among Subjects in Choosing Trigram Labels for Random Shapes" (Vol. 83, No. 1, p. 68) should read:

BERNHARD KEMPLER, *Georgia State University*,
and HERBERT KAYE, *Emory University*

THOMAS VERNER MOORE: 1877–1969

On June 5, 1969, Thomas Verner Moore, one of Wilhelm Wundt's last living students, died in his ninety-second year at the Cartuja de Miraflores, Burgos, Spain. Although he had been living over twenty years in the seclusion of a Carthusian monastery, his death was widely noted in the Spanish press. He had gone to Spain in 1947, after retiring as head of the department of psychology and psychiatry at Catholic University, to give a series of lectures on psychiatry at the University of Madrid, and to remain in Spain as a Carthusian monk.

This writer tried to visit him in Burgos in March, 1968, but was told that since it was Lent, the period of *El gran silencio,* he could see no one. He was reported to be well, but feeble and keeping mostly to his room. He had come to the United States in the mid-1950s to establish the first American Carthusian monastery, in Vermont; at that time he paid a brief visit to Catholic University, lecturing to the students of psychology. As a Carthusian, he had continued his interests in psychology, but it was a psychology transmuted into something beyond science: his last two books, published from Burgos, were *Heroic Sanctity and Insanity* and *The Life of Man with God.* But he never lost touch with empirical methods: as a Carthusian, he did a factor-analytical study of the states of mystical experience, which is perhaps the ultimate attempt to combine the empirical with the intangible.

Father Moore was born in Louisville, Kentucky, on October 22, 1877. He received, from Catholic University in 1903, the first doctorate in psychology granted in the District of Columbia. This he published as a monograph in 1904 with the title "A Study in Reaction Time and Movement." He was then sent by Edward A. Pace, a student of Wundt and the founder of the psychology laboratory at Catholic University in 1904, to Leipzig to study under the master. Given only a short time to live because of tuberculosis, he returned to the United States to finish and publish at the University of California the psychological research begun under Wundt, *The Process of Abstraction* (1910).

In 1910 he was back on the staff at Catholic University. After a visit to Witmer's pioneer clinic at the University of Pennsylvania, he decided he should have a medical degree so that he could work with disturbed children. So he went to Munich to study medicine. These studies were interrupted by World War I. He returned to the United States to finish his medical studies at Johns Hopkins, taking his M.D. there in 1915. He then opened a clinic for children at Providence Hospital, Washington, and introduced the first clinical course at Catholic University, a course that was titled A Clinic for the Examination of Defective Children. Shortly, however, he was back in Europe as a psychiatrist with the American Expeditionary Forces.

After World War I, he succeeded Dr. Pace as head of the department, a post he held for almost thirty years until his retirement from the University in 1947. In 1926, he established Saint Gertrude's School of Arts and Crafts for retarded children in Washington. In 1937, his clinic at Providence Hospital was moved to the campus of Catholic University, where it became The Child Center, a charter member of the American Association of Psychiatric Clinics for Children (supported for a number of years by the Rockefeller Foundation).

Father Moore's pioneer contributions to psychology were fourfold. He was a pioneer in integrating theoretical concepts, experimental methods, and clinical practice. He was a pioneer in applying the data from abnormal psychology to the understanding of the functioning of the normal personality. His *Dynamic Psychology* (1924) was one of the early books to use that title. He was a pioneer in making psychology respectable in Catholic circles: he wrote appreciatively of Freud, Jung, and Adler when these men were almost unknown to Catholics (and to most others, too), and if known, practically anathema. He was a pioneer in his interest in human behavior in *all* its aspects, from the deeply disturbed psychotic to "the life of man with God." In many ways he was ahead of his times: in the 1930s he and his students were doing research on the influences of drugs on psychic functioning; he established a reading clinic as a part of his child-guidance center; he was an early student of the mathematical analysis of human abilities.

He was a quiet, genial, friendly man. Empiricist and practictioner, he was also the detached, visionary, and often absentminded scholar. As a clinician and psychiatrist, he knew clearly and pain-fully the problems of his fellow man; as a philosopher and dedicated man of God, he knew what man could become. His life was devoted to the task of helping man to a fuller and richer life, here and hereafter.

JOHN W. STAFFORD, C.S.V., *The Clerics of Saint Viator*
Arlington Heights, Illinois

BOOK REVIEWS

Edited by JULIAN HOCHBERG,
Columbia University

THE MEASUREMENT AND PREDICTION OF JUDGMENT AND CHOICE. By R. DARRELL BOCK and LYLE V. JONES. San Francisco: Holden-Day, 1968. Pp. 370. $13.75.

This book is an intensive elaboration of Thurstonian methods for psychological scaling. Following a brief introductory chapter, there are four chapters devoted to the measurement of absolute and difference thresholds. Chapters 2 and 3 discuss the statistical and computational problems in fitting normal ogives (or other two-parameter classes of cumulative distributions) to psychometric functions obtained by the method of constant stimuli. Chapter 4 deals with certain variants of the constant-stimuli method that are useful in studies of absolute thresholds. Chapter 5 then gives a brief statistical treatment of Weber's law.

The remaining chapters (Chapters 6–9) treat the measurement of 'affective values' by analysis of confusion probabilities in the methods of paired comparisons, ranking, or category rating, as well as various applications once the affective values are obtained. Such applications include linear models relating affective values to quantitative or qualitative characteristics of the stimulus objects, the use of affective values to predict choices from among more than two alternatives, and the prediction of choices between compound stimuli by combining affective values of the elemental stimuli. Appendix I then contains an unusual and valuable set of tables.

The book is quite clearly written and should be easy reading for anyone who has had enough mathematical statistics to be comfortable with maximum-likelihood and related estimation procedures. But I have very mixed feelings about the value of this book. The authors treat some things extremely well. Most notable are the many detailed examples, the consistent attention to proper estimation of confidence intervals, and the beautiful section on incomplete paired comparisons. Other valuable material includes treatment of the Newton-Raphson method for solving nonlinear estimation equations, with the relation of this method to probit analysis; discussion of bias in small-sample point estimates; the triangle test (oddity detection) for absolute thresholds; and inclusion of individual components in models for comparative judgments.

On the other hand, a number of things are done rather poorly. A few of the mathematical developments are unnecessarily hard to follow, either because of unfortunate notation, misprints in formulas, or logical non sequiturs. These will not hamper a reader who knows mathematical statistics, but they may cause difficulties for someone trying to teach himself these methods with this book.

The shoddy treatment of psychological theory is more serious. Two annoying, though unimportant, examples are found in Chapter 1: the quaint distinction between 'judgment' and 'choice' (the correct basis for a distinction is what the experimenter assumes—that the subject's preferences are known or that the stimuli are perfectly discriminable—not what the subject is trying to do) and the simplistic view of model evaluation (that a model with fewer parameters showing a better fit is always superior). More important is the fact that the book ignores modern psychophysical thinking. The problem starts on page 3, where the distinction between theories involving variable thresholds and theories involving variable stimulus effects is muddled, and it extends through the first five chapters. Indeed, there is not a single reference in the text or the bibliography to the literature on signal detection, despite numerous interconnections in Chapters 2–5 and Chapter 8.

The authors of this book repeatedly recommend counterbalancing spatial or temporal order of stimuli and combining of the data from the counterbalanced orders, to eliminate errors due to 'order effects.' In justification they offer a simple symmetric order-effect model (p. 17), but this model is not seen again and is inconsistent with ones developed later in the book. Insofar as these later models are really appropriate, they contradict the recommendation to combine the data from counterbalanced orders. No mention is made of Luce's work on order effects nor of relevant theoretical issues such as the separation of response biases from real effects of spatial or temporal order. The treatment of Weber's law is similarly insulated from modern thinking on the subject. The authors use a form of Weber's law that cannot possibly hold for low stimulus values; and they repeat Fechner's totally incorrect derivation of the logarithmic transformation. (Their method yields the correct answer in special cases by accident; their Equation 5.9 does not yield even a correct first approximation in general.)

The authors ignore modern work on probabilistic choice theory almost as much as they ignore psychophysical theory, although the consequences my be slightly less serious for the applied purposes of the book. The list of specific criticisms could be extended considerably.

The significance of these failures in the area of psychological theory must be evaluated in the light of the authors' expressed goals, however. For example, one of the major objectives of Chapters 6–9 is the prediction of choice frequencies of each of several alternatives by members of a specified consumer population. Suppose, for instance, that a set of choice alternatives varies on five factors and that each factor has ten discriminable levels. This yields $10^5 = 100{,}000$ potential stimuli. Suppose further that one wishes to develop a general formula to predict approximate choice frequencies when an arbitrary subset of n alternatives is offered to the consumer population. This book gives some practical methods that may work. For example, one might test levels 2, 5, and 8 on each factor, for a total of $3^5 = 243$ stimuli. A random sample from the consumer population might rate these 243 stimuli singly (Chapter 8) or might rank subsets of them in a balanced incomplete-block design (Chapter 7). Using an appropriate model (Chapters

6-8), affective values for the 243 stimuli could be estimated. Then a linear prediction formula (Chapter 7) for affective values, in terms of quantitative measures of factor levels, could be developed, permitting estimation of affective values for any of the original 100,000 stimuli. These estimated affective values would, in turn, allow predictions of the desired choice frequencies in subsets of *n* stimuli (Chapter 9). One could also test goodness of fit of both the model relating affective values to choice probabilities and the linear model relating affective values to factor levels. Assuming a reasonable fit of these models, one could calculate confidence intervals for the predicted choice frequencies. Thus, the authors' major objectives in Chapters 6–9 are quite well fulfilled. Indeed, the entire book can be judged more favorably if it is considered as a contribution to applied statistics rather than to psychology.

When the Thurstonian paradigm was developed, there was little competing theory in the areas of psychophysics and preferential choice. In the past 25 years, however, substantial bodies of theory and of experimental results have accumulated in both areas; in psychophysics especially, a conceptual revolution has occurred. The Thurstonian paradigm seems to have proceeded along its own insulated pathways, ignoring these developments. In so doing, it has lost much of its status as psychological theory and has become, rather, a methodology for prediction of event frequencies by extrapolation from small samples and restricted situations. It has become, in other words, a branch of applied statistics.

How useful is this body of applied statistics to psychologists? In the area of psychophysics, where conceptual growth has been greatest, I am afraid that the Thurstonian methodology may be dangerous, not so much because it generates wrong predictions in limited extrapolations, but because it doesn't provide the right framework for understanding what is going on. In the area of consumer choice, no one understands what's going on anyway, and the use of this methodology may be quite satisfactory for the time being. For those concerned with predictions of consumer choice, this book should be valuable.

David Krantz, *University of Michigan*

THE TASK OF GESTALT PSYCHOLOGY. By Wolfgang Köhler. Princeton, N. J.: Princeton University Press, 1969. Pp. vii, 164. $6.50.

In November of 1966, only a few months before his death, Wolfgang Köhler was called upon to give the Herbert S. Langfeld Memorial Lecture Series at Princeton University. These four lectures have been reprinted as *The Task of Gestalt Psychology,* although on the basis of content, the title of this volume could just as easily have been "An Overview of Gestalt Psychology."

The fact that these lectures were delivered to a mixed group of listeners, not all of whom were specialists in psychology, is both a point of weakness and of strength. It is a weakness because we are denied another look at Köhler's incisive approach to theoretical problems—an approach that tends to include not merely demonstration in the traditional manner but also careful experimentation. It is a strength because, freed of the necessity of specifying the details a colleague in the field would require of him, Köhler can survey the broad expanse of Gestalt theory and attempt to show the outlines of a unified approach that has been applied in many areas of psychology.

The opening lecture is familiar to us all. It is the traditional Gestalt attack on the atomistic and analytic approach in psychology. Beginning with an analysis of the theoretical importance of stroboscopic movement, it moves through the concept of form quality and ends with a discussion of the general nature of interaction. This chapter could have as easily been produced by Wertheimer or Koffka. It is the battle cry of the Gestalt school of thought. The second lecture is quite different in tone and style. It is a philosophical justification of the concept of isomorphism. Köhler begins with an interesting and insightful discussion of Descartes' mechanistic view of the operation of the nervous system, and ends this section with a discussion of evolution. This is not psychology; it is, however, Köhler at his best.

The third lecture is the most data-oriented of the entire volume. It begins with a discussion of cortical brain fields and reaches its peak with a presentation of some of the Swarthmore data on electrical brain correlates of visual stimulation. Köhler does not stop to deal with the objections of critics such as Lashley or Sperry; in fact, he does not even recognize their existence in this phase of his work. In this section, one is given a brief glimpse of the fact that besides being an excellent theorist, Köhler also was a competent and clever experimentalist. At this point, there is a sudden transition within the chapter and the discussion shifts to the field of memory. Here Köhler attempts to bring memory under control of the same organizing principles used in perception, and he includes a rather detailed discussion of the von Restorff effect. The fourth lecture then deals with relational learning and the problem of insight. It is primarily based on the data taken from the Teneriffe station —data which were the basis for *The Mentality of Apes*—but it is updated with some interesting examples of mathematical insight. This final section closes rather abruptly without a general summing-up of the central theme of the theory, which summing-up might have been helpful.

Perhaps the weakest part of the book is the preface by Carroll C. Pratt. Though it provides some useful biographical data, it spends nearly one-third of its length on the problem of aesthetics, which Köhler does not discuss at all. Pratt also fails to analyze the impact of Gestalt theory on contemporary work. Given the fact that this is the final work of one of the theory's originators, such an analysis seems appropriate. But overall this volume has been done well. It is written in the persuasive and eloquent style that was Köhler's trademark, and it provides a panoramic view of the breadth and scope of Gestalt theory. It is not a book for the specialist in search of specific data or analysis; it is, rather, an interesting, warm, and intelligent overview of one of the most important theoretical systems to be introduced into American psychology in the last 50 years.

Stanley Coren, *New School for Social Research*

RECOGNIZING PATTERNS: STUDIES IN LIVING AND AUTOMATIC SYSTEMS. Edited by Paul A. Kolers and Murray Eden. Cambridge, Mass.: The M.I.T. Press, 1968. Pp. xi, 237. $11.00.

This interesting introduction to pattern perception is defined by Kolers' overview chapter on animals' pattern perception, Eden's summary chapter, which provides a general discussion of the machine problem, and a collection of contributed chapters on both topics sandwiched in between. The contributed chapters cover living systems from the points of view of speech production and perception (Samuel Jay Keyser and Morris Halle), visual neurophysiology (Shin-Ho Chung), and peripheral auditory mechanisms (William M. Siebert). Machine recognition is discussed in terms of recognition of handwriting (Murray Eden), of the development of an experimental reading machine for the blind (Samuel J. Mason and Jon K. Clemens), and of the development of programs whereby contextual understanding provides a man-machine link (Joseph Weizenbaum).

The major strength of the book is its breadth, which is provided both in the definition of what pattern is and by the diversity of approaches to the study of pattern recognition (a diversity represented by the various contributed chapters). Psychologists will find Eden's introductory chapter an excellent overview of pattern perception from *their* point of view. The major weakness of the book is that the hoped-for synthesis does not occur. One is led to believe by the very nature of such a book that the general principles which allowed one to traverse the route between the first and last chapters would be made explicit. The emphasis on empirical results, post hoc evaluation, and a diversity of approaches represents, however, the state of the art. From reading this book one might conclude, as one of the authors suggests, "that all that is shared by those who would build seeing machines and those who investigate existing biological perceivers is a respect and admiration for any system that works" (p. 99).

DONALD R. BROWN, *Purdue University*

CLASSICAL CONDITIONING. By ROBERT S. BEECROFT. Goleta, Calif.: Psychonomic Press. 1966. Pp. ii, 198. $5.25.

In reviewing the American literature on classical conditioning, Beecroft has performed a very useful service. He has organized and summarized over 400 articles on the classical conditioning of a variety of response systems. The coverage of conditioning of the eyelid, heart-rate, pupillary, skeletal-motor, EEG, and galvanic skin responses is fairly extensive, and there is some mention of nictitating membrane, vasomotor, and salivary responses. The author has included historical and methodological accounts of the general problems encountered in research on classical conditioning, and he has discussed some of the special problems associated with particular response systems. He has also summarized the literature on simple factors that influence the conditioning process and on such other traditional phenomena as extinction, generalization, and compound conditioning. The book concludes with a brief summary of some theories of classical conditioning and of other, broader theories that have been based on classical-conditioning models.

The readability of the text suffers from the extremely condensed nature of the exposition, the lack of any figures, and an almost complete absence of actual data. Also, the author has included a large number of strong statements of opinion; these are not always supported by data, and they tend to be simplistic and somewhat misleading. Although there is little coverage of literature beyond 1963 or for other than American research, the book should be useful to advanced students as a fairly comprehensive introduction to the area of classical conditioning, and it is a good source of references and summaries for professional workers.

DAVID C. RASKIN, *University of Utah*

DEVELOPMENT IN HUMAN LEARNING. Edited by E. A. LUNZER and J. F. MORRIS. New York: American Elsevier, 1968. Pp. xiv, 487. $10.50.

The volume under review is the middle part of a trilogy with the general title *Development in Learning*. Volumes 1 and 3, which I have not seen, are *The Regulation of Behaviour,* which treats issues in the study of learning, and *Contexts of Education,* dealing with individual differences and social factors in education.

Volume 2, the book with which we are concerned here, consists of 14 chapters by 13 contributors: "Perception and Perceptual Learning," by M. D. Vernon (one of the book's irritating stylistic oddities is the failure to give authors' first names); "Social Learning and Identification," by H. R. Schaffer; "Experimental Approaches to Concept Development in Young Children," by Marion Blank (an exception whose given name is spelled out and who is also the only non-British contributor); "Skill," by G. F. Reed (discussing not so much 'skills' as 'skillfulness,' or mastery); "Language and Mental Development," by M. M. Lewis (after he has spoken repeatedly of "orectic factors," Lewis finally identifies these as "personality"; Lewis refers to me as an "anti-Skinnerian," whereas, being anti–almost everybody, I tend to think of myself as a Churchian); "The Experimental Psychology of Grammar and Its Relevance to Children's Language Learning," by P. Herriot; "Reading and Learning to Read," by G. R. Roberts and E. A. Lunzer; "Systematization of Thought," by K. Lovell; "Formal Reasoning," by E. A. Lunzer; "Conceptual Learning and Explainer Thinking," by E. A. Peel; "Social Learning and Perspectives in Adolescence," by J. F. Morris; "Programmed Instruction," by H. Kay; "Learning and Development—Schematic Learning in Mathematics," by R. R. Skemp (as a mathematical semi-literate, I found Skemp's paper entertaining and informative, but I am not sure how it relates to learning in general; there are overtones of the value of 'mental discipline,' but this may have been my own projection); and "Psychology and the Teacher," by E. A. Lunzer (a pulling-together of the book and of the three volumes).

I have listed the contents in detail so that I could make there the few remarks I wanted to make about individual chapters rather than about the book as a whole. I wanted to give the reader an overview of its contents and, in so doing, to pose a concept-formation task: what does the title of this book mean? As nearly as I can make it out, the book should have been called "Learning in Human Development," since this seems to be the unifying strand.

This book has several American counterparts, notably the two-volume Hoffman and Hoffman *Review of Child Development Research* (New York: Russell Sage Foundation, 1964 and 1966) and the forthcoming Mussen revision of Carmichael's *Manual of Child Psychology* (New York: Wiley, 1954). While the coverage is much the same, the organization of topics is rather different; there is much more attention in the British volume to the British literature, and the perspective is indefinably but tangibly British. The conributions to the Lunzer and Morris volume are consistently thorough, scholarly, and—with a few happy exceptions that I prefer not to single out—about as badly written as the American psychological literature. At first, one finds oneself rewriting as one reads, which is probably an excellent way of ensuring attention and comprehension; but this process soon becomes tedious, and the reader finds himself sliding through whole paragraphs with zero retention, so that he is forced to renegotiate the same passages in search of enlightenment. The book is haunted by two presences, one living and one phantom. The living presence is Jean Piaget. The extent to which he dominates these pages can be gleaned from the number of times he is cited in the index: 34 mentions, closely followed by 32 for his collaborator Inhelder (there is considerable overlap), as compared with 15 references to Skinner, 13 each to Bartlett and Vigotsky, 12 to Bruner, 10 to Lovell (one of the authors, whose own chapter of course figures among the mentions), 9 to E. J. Gibson, 8 to H. Kay (also a contributor), 7 each to J. Annett, Broadbent, Harlow, E. A. Peel (another author), and Wohlwill, 6 each to I. H. Anderson, Chomsky, W. F. Dearborn, Erikson, Freud, Gesell, E. Kaplan, H. W. Reese, and H. Stevenson, and way down the totem pole, 5 each to A. Alpert, Bandura, R. Brown, C. C. Fries, Fromm, J. Kagan (mislabeled in the index as L. Kagan), Lorenz, Werner, and Wertheimer. It must be said that Piaget is treated here less reverently than by some of his American disciples, but he is undoubtedly the key character. The phantom presence is information theory, which author after author discovers as the harbinger of miracles to come, apparently unaware that they are reenacting an American drama of 20 years ago, in the final act of which the decomposed remains of information theory were gently laid to rest, leaving behind only a few metaphorical mementos (feedback, channel, information overload, and the rest). However many analogies one can find between information concepts and human processes, information-theoretical analysis leads in the end not to psychological elements but to computer programs. Computer simulation of behavior may be an amusing intellectual game, but I find it hard to believe that anyone nowadays takes it seriously as an approach to understanding the human condition.

The jacket copy states that this book is "designed to serve as a course textbook for students in Departments of Education who are preparing for higher diplomas and degrees." I cannot assess how well it serves this function, but it seems to me that the approach is open to question. The book, like its American counterparts, may collapse between two stools. It lacks the virtues of either the true textbook, which is woven together, no matter how idiosyncratically, out of a single vision, or the book of readings, which puts the student in direct touch with some of the original documents of a field. Collections such as this can serve mainly as useful reference works until, in a few years, they have gone completely out of date.

JOSEPH CHURCH, *Graduate Center, City University of New York*

CHEMICALLY ALTERED STATES OF CONSCIOUSNESS: NINE REVIEWS

The Psychopharmacology of the Norman Human, edited by Wayne O. Evans and Nathan Kline (Springfield, Ill.: Charles C Thomas, 1966), is a 252-page book and one of the most recent collections of research papers on the effects of psychoactive drugs. In addition to containing some useful and interesting results, it provides evidence that this research area is maturing. Witchcraft, mythology, and medical dogmatism are being replaced by objective experimentation and analysis.

The book contains nine chapters by twenty contributors. Comments on four of these chapters follow. "The Effect of Antianxiety Tranquilizers on the Behavior of Normal Persons" is by F. M. Berger and J. Potterfield, who found that minor tranquilizers such as meprobamate (Miltown) when used in the conventional clinical dosages do not alter the behavior of emotionally stable people but tend to "normalize" the behavior of unstable persons. This result is in contrast to the effects of older sedative-hypnotics such as the barbiturates, which alter behavior independent of emotional stability. "The Effects of Amphetamine and Chlorpromazine on Cognitive Skills and Feelings in Normal Adult Males" is by E. K. Davis, W. O. Evans, and J. S. Gillis. They compared *d*-amphetamine (Dexedrine) (15 mg.), chlorpromazine (Thorazine) (75 mg.) and a placebo, using a battery of cognitive tests on 90 adult males. The effects were not large because the dosages were not large. (Teenagers often 'pop' 50 mg. of Dexedrine, and dosages of Thorazine in psychotherapy run as high as 800 mg.)

"Studies of Psychoactive Drug Effects on Nonpsychiatric Patients: Measurement of Affective and Cognitive Changes by Content Analysis of Speech" is by L. A. Gottschalk, G. C. Gleser, W. N. Stone, and R. L. Kunkel. Stimulants, sedatives, major and minor tranquilizers, and three psychedelics (LSD, ditran, and psilocybin) were used in a series of experiments from which content analyses of speech were made. The instruments used were not sharp and precise, but that didn't matter much: the effects observed were large and therefore very interesting. Anxiety scores were reduced by amobarbital and chlordiazepoxide, two rather strong depressants. Achievement and energy-level scores were increased by amphetamines. The three psychedelic drugs did not alter anxiety and hostility scores but increased "cognitive and intellectual impairment scores." It was also found that set or suggestion regarding the drug effects were about as effective as the drugs themselves. The problems were attacked head-on and so the results were rewarding—unlike those of many studies in which subtle effects were measured with precision but from which one learns very little.

"The Attenuation of the External Environment and the LSD Effect" is by Allan Edwards and Sidney Cohen. Thirteen LSD-naïve subjects were tested under four conditions: with 125 μg. of LSD or a placebo, under sensory deprivation or sensory normalcy. Under sensory normalcy, all subjects correctly reported the presence or absence of LSD. But in the LSD sensory-deprivation condition, seven subjects reported no LSD taken and the other six reported that they had had a much reduced dose of LSD. The investigators realized in advance that sensory deprivation is never complete in such experiments. This and other evidence support the view that complete sensory deprivation would fully counteract the LSD effect. The conclusion seems clear: LSD manipulates sensory inputs; it does not manufacture them. This is an unpretentious small study conducted by an expert in experimental design and a distinguished psychopharmacologist—which seems to make a good research team in the field of psychoactive drugs.

This first review raises a much larger issue, which deserves at least brief discussion here. Psychologists should be more interested in psychoactive drugs such as amphetamines, barbiturates, heroin, and LSD for at least two good reasons: First, the drugs' effects are becoming more and more socially relevant, and much scientific work is needed. Second, these chemical agents are promising research tools in developing better theories of cognitive processes, indeed, of consciousness itself. A relatively small number of psychologists have been aware of this for a long time, but the possible risks to human health as well as governmental restrictions (some of them unnecessary) have inhibited psychological investigations and restricted them to specialized medical settings. A few psychologists such as Timothy Leary have become crusaders or cultists and have thus isolated themselves from the mainstream of psychology. But by and large, psychologists have avoided working with the chemical mind-changers for one reason or another.

A very powerful mind-altering drug, ethyl alcohol, has been America's national 'recreational drug' for a long time, but in recent years it has acquired a competitor: marihuana. The National Institutes of Health estimated that 12,000,000 Americans—most of them young—have tried this weed. That figure is still small compared to the 80,000,000 drinkers and 6,000,000 alcoholics in the country. The use of heroin and other opiates has been on the decline for decades (their publicity, however, has been increasing); but amphetamines and other stimulants, and barbiturates and other sedatives, are being used in heavy dosages more every year. Finally, the psychedelics, or hallucinogens, such as LSD, mescaline, and psilocybin are now being used for the first time in history as recreational drugs. (Their use for religious or therapeutic goals dates back several centuries.)

From the point of view of psychological research, the psychedelic drugs are the most interesting, because they can produce striking new phenomena. The opiates and related drugs decrease a person's sensitivity to external and internal stimuli. The amphetamines seem to speed up mental and physical processes. The barbiturates and tranquilizers, like alcohol, slow down those processes. But the psychedelics seem to 'expand consciousness,' as the LSD devotees say, and they certainly cause perceptual distortions and illusions. Sometimes they appear to cause temporary or permanent psychoses; more likely they tend to facilitate the development or expression of psychoses.

My interest in psychoactive drugs became more than casual less than a year ago, when a request to teach an undergraduate seminar on the topic led me to read a large part of the existing literature. (I soon discovered that many of my students were better read than I.) Because I had no professional guide, I did a good deal of stumbling about. The following is intended to save others some time and effort if they want to become familiar with the field. The major books are briefly described below.

The Pharmacological Basis of Therapeutics, edited by L. S. Goodman and A. Gilman (3d ed.; New York: Macmillan, 1965), is a 1,785-page compendium, written by 42 leading authorities, on all kinds of drugs used by the medical profession. Section II, "Drugs Acting on the Central Nervous System," contains most of the material on psychoactive drugs. Chapters 9 and 10, by Seth K. Sharpless, deal with hypnotics and sedatives—the barbiturates and similar depressants. Chapter 15, by Jerome H. Jaffe, is concerned with the opiates and other narcotics (as medically defined, not legally). The amphetamines are in Section IV, Chapter 24, on "Sympathomimetic Drugs" by Ian R. Innes and Mark Nickerson; and cocaine, also a stimulant, is treated by J. Murdoch Ritchie, Peter J. Cohen, and Robert D. Dripps in Section III on "Local Anesthetics" because this is the sole medical use of cocaine today. The psychedelics are covered almost incidentally in Chapter 12, by Murray E. Jarvick, on "Drugs Used in the Treatment of Psychiatric Disorders." Indeed, the medical use of the psychedelics has been almost incidental. Fortunately, a separate medical book on psychedelics is available (Hoffer and Osmond, see below). Although the organization of Goodman and Gilman leaves much to be desired by the psychologist, it is by far the most complete orthodox medical book on drugs; a fine index makes it even more useful. The only major weakness of the book is a reflection on the state of pharmacology: it lacks clean and complete experimental data using reliable and valid psychological measures. Furthermore, modesty and caution are lacking; often a single, poor experiment is the basis of a 'statement of medical fact.' But if you want to know what is known, Goodman and Gilman is the best source.

The Hallucinogens, by A. Hoffer and H. Osmond (New York: Academic Press, 1967), is by far the most complete scientific treatment of psychedelic drugs. It contains most of the historical, anthropological, experimental, clinical, chemical, and physiological knowledge on all known drugs in this class, as well as the major theoretical attempts at explaining their effects. The drugs are classified chemically because no other way has been seriously proposed. Cannabis is not included because the authors apparently do not consider it in the same category (nor do I).

Drugs from A to Z: A Dictionary, by R. R. Lingeman (New York: McGraw-Hill, 1969), is an amazingly complete and accurate book in dictionary form, written by a journalist. In paperback, it sells for $2.95 and so is a 'must.' It includes slang or 'hip' terms as well. It is impressive that a layman can be as knowledgeable as Lingeman clearly is. This should also be encouraging to students.

Junkie, by W. S. Burroughs (New York: Grove Press, 1957), is a partial autobiography by one of our most talented creative writers. It describes his experience as a narcotics addict for 15 years. His book, *Naked Lunch* (New York: Grove Press, 1959), which is more difficult to read, describes the drug-induced experiences of a brilliant man and writer. It has an appendix describing the effects of the major 'recreational' drugs—an appendix that might teach a lot to physicians.

The Marihuana Papers, edited by D. Solomon (New York: Bobbs-Merrill, 1966), is the most important, though a bit dated, collection of societal, literary, and scientific papers on the milder form of cannabis, including the classic LaGuardia report, Howard Becker's "Becoming a Marijuana User," and a Burroughs essay. The orientation is definitely pro-marihuana, but therein may lie its value. It comes in paperback for only $1.50.

Pot: A Handbook of Marihuana is by J. Rosevear. (New Hyde Park, N. Y.: University Books, 1967). Despite its subtitle, it furnishes the scientifically inclined an abundance of information on nature's source of tetrahydrocannabinol and probably other related psychoactive drugs. (I agree with Burroughs that cannabis does not belong in any of the four major categories of psychoactive drugs: stimulants, depressants, narcotics, and psychedelics. Rather, cannabis is a sensitizer, enhancer, or facilitator of moods and effects of other drugs.)

The Varities of Psychedelic Experience is by R. E. L. Masters and J. Houston (New York: Dell Books, 1966). Many books and articles have been written about the horrors of LSD and about the beneficial effects of the drug. Among those books sympathetic to LSD use, the best I have seen is this one. The authors present lucid descriptions obtained from first-hand observations on over 200 drug sessions and from interviews with another 200 persons who had had psychedelic drug experiences. The book provides an account of the 'richness' of the phenomena. The authors are clearly opposed to current legal and social restrictions on LSD use but are equally opposed to its unrestricted use.

Physicians' Desk Reference to Pharmaceutical Specialties and Biologicals (23d ed.; Oradell, N. J.: Medical Economics, Inc., 1969), known in the trade as the 'PDR,' is a compilation of manufacturers' information on their drug products. Over 2,600 drug products are cross-indexed in various ways. The main value of the PDR is in identifying a legal drug from partial information such as trade name, chemical name, manufacturer, or appearance (color photographs are included). The PDR is given free to all medical practitioners but is otherwise not generally available. My own experience suggests that it can be obtained by researchers and teachers, however.

ROBERT R. BUSH, *Columbia University*

FUNDAMENTALS OF SKILL. By A. T. WELFORD. New York: Barnes & Noble, 1968. Pp. 426. $11.25.

Those who have tried to keep abreast of research and thinking in British experimental psychology will be grateful for this book because it is a good statement of the topics that interest so many British experimentalists and which they subsume under the rubric of skills. The word 'skills' connotes motor behavior to many psychologists, but Welford dismisses this narrow definition and says that skills are "concerned with all the factors which go to make up a competent, expert, rapid and accurate performance. Skill in this sense thus attaches, to a greater or less extent, to *any* performance and is not limited to manual operations but covers a wide range of mental activities as well" (pp. 12–13). Such an all-embracing definition can cover the concerns of most psychologists, and so robs 'skill' of the precision required to make it a scientifically useful term, but the contemporary history of experimental psychology in Britain makes Welford's definition understandable. Sir Frederic Bartlett was a prime mover of British experimental psychology and his involvement in human-factors work during World War II stimulated an interest in man-machine relations that left its mark on Bartlett and the psychologists he trained. Bartlett's progeny now exert a prominent influence on British experimental psychology, and man's behavior in work settings is the implicit frame of reference which drives the interest in skills that many of them have. Welford's definition of skills makes sense, given these determinants in his scientific culture. Instead of being solely the study of motor behavior, the study of skills becomes a concern with all the variables that make up the complex behavior of factory workers who operate lathes, of pilots, or of radar operators. Yet the study of skills is not necessarily applied psychology. Perhaps from Bartlett comes a strong urge toward basic research and toward the point of view that the laboratory and the field enrich each other. Those who tease apart the variables in the laboratory will someday have their say in the field, and they choose research problems that stand a better-than-even chance of eventually bearing on practical work situations. Welford's book should be read with this scientific attitude in mind; it reflects an ease of movement from field to laboratory and back again, and lacks the tendency toward compartmentalization of basic and applied research that characterizes psychology in the United States.

As might be expected, Welford's theorizing has a strong engineering flavor, and he reveals a conviction that engineering theory and psychological theory are interdependent and interchangeable when he says, "It is a not unfair test of a psychological theory to challenge its author to design a machine that would behave in the same way as his theory would predict for a human being" (p. 25). This bias is not likely to catch on in theoretical psychology or to stir philosophers of science, but it is enough to justify Welford's use of signal-detection, information, and servo theories in his treatment of behavior. His use of servo theory is by analogy, which is common, but his applications of signal-detection theory and information theory are often mathematical. Despite theoretical inclinations, Welford cannot escape the fact that much of the psychology of human behavior today is empirical, and so much of the book is empirical. The text is almost entirely about behavior, with small patrols occasionally run into physiological territory.

For the most part the chapters reflect the mechanisms that Welford sees as the ingredients of skills. Chapter 1 gives the conceptual framework and a block-diagram schema of the mechanisms to be elaborated in subsequent chapters. Chapter 2 is about simple decisions, such as saying whether a stimulus occurred or not in a psychophysics experiment; and it is mostly an exposition of signal-detection theory, rather than a presentation of data on stimulus discrimination. Chapter 3 is primarily about choice-reaction time, and it pivots around Hick's use of information theory to describe reaction time as a function of number of stimulus alternatives. There is no attempt to give a complete coverage of reaction time or to include such topics as simple reaction time, foreperiods, or stimulus properties, but this is an expected focus for an information theorist who needs multiple events and their probabilities on which to operate. Chapter 4 is about doing two or more things at once or in rapid succession, and it is Welford's prize chapter, perhaps because the British have been in the forefront of this research. The heading "Single-Channel Operation" reflects the British conviction that man is a limited data-processor and can handle only one event at a time, making him an "intermittent-correction servo" because he must respond, process, respond, process, and so on. Chapter 5 represents an effort to clarify motor movement, and it centers on Fitts' work on movement time as a function of target width and distance. Chapter 6 opens with a conventional treatment of perception and the meaningful processing of complex stimulus displays; it ends with a satisfying account of the literature on control-display linkages, saying that the perceptual interpretation of displays is only part of the behavioral battle.

Chapter 7 is a review of short-term memory; it gives too much space to studies of memory span and dichotic listening and not enough to the other data and ideas on short-term memory that have deluged us in recent years. There is a refreshing challenge of the trace-decay theory of forgetting, which is surprising because trace decay has been orthodoxy in British thinking for some time. Regrettably, Chapter 7 lacks a sister chapter on long-term memory. Loading is the topic of Chapter 8—a long chapter because it spans a range from the work overload that produces fatigue to the underload of infrequent signals that creates a need for vigilance. The section on fatigue has interesting ideas on the locus (central versus peripheral) of fatigue, but it is deficient because it bypasses the American research on work inhibition. Chapter 9 on acquisition of skill is sound enough as far as it goes, but only one chapter in ten on acquisition gives learning a minor role in Welford's scheme of things. That learning should be an unauspicious citizen in a land where engineering theories reign is not surprising. The final chapter, Chapter 10, on personality, motivation, and social relationships, is so outside the mainstream of the book as to be an anomaly. Perhaps the chapter represents Welford's attempt to broaden the scope of inquiry for skills, which is commendable, but it comes through as an afterthought and one cannot escape the feeling that it should have been left out.

None of these criticisms should detract from the conclusion that this is a scholarly book written by a man who knows what he is talking about. The 1,000 or so references, integrated and interpreted, show Welford's depth of penetration, which is also present in his clear prose. The book invites comparison with two other recent books on skills (E. A. Bilodeau, *Acquisition of Skill*, 1966; E. A. Bilodeau and I. McD. Bilodeau, *Principles of Skill Acquisition*, 1969). Happily, the redundancy is minor. Any psychologist interested in the complexities of human performance, whether he be of 'basic' or 'applied persuasion,' will want all three of these books on his shelf.

JACK A. ADAMS, *University of Illinois at Urbana-Champaign*

FOUNDATIONS OF EXPERIMENTAL RESEARCH. By ROBERT PLUTCHIK. New York: Harper & Row, 1968. Pp. xiii, 290. Paperback, $3.95.

This small volume is essentially a textbook on the logic of research. The fact that the author is a psychologist is detectable but not intrusive. The writing is succinct but the coverage is general. The beginning student will learn the basic reasons underlying various kinds of experimental design, laboratory methodology, and statistical evaluation, but he may not be in position to do any research, because of his lack of contact with any real content or with any experience in the form of exercises in which the logic has been employed. The book is written as a series of explanations of why researchers proceed as they do, of why they use this or another measure or statistical test—without taking into account the student's interests. It appears to be more an aid for instructors who are teaching elementary experimental or statistics courses and may not be quite sure of what they are about. In the latter circumstance it could be quite useful to instructors whose basic preparation has been hasty or inadequate.

The reviewer felt a lack of contact with psychology in any historical sense —of any relation of the methodology to psychological problems. There is a brief chapter on psychophysics, but the student will not be able to grasp the concepts involved in Weber's law from its treatment there. One gets the impression that psychological methodology developed yesterday, full-blown, ready to apply—without any appreciation of where and to what it does apply. In sum, much of the material is competent and valuable, but there is no question that the text needs a teacher, a program of exercises, and considerable support from additional readings. It is clearly a supplement and, in the right hands, could be very helpful.

B. R. Bugelski, *State Uiniversity of New York at Buffalo*

METHODS FOR EXPERIMENTAL SOCIAL INNOVATION. By George W. Fairweather. New York: Wiley, 1967. Pp. x, 250. $7.95.

This book is both a textbook and a book of advocacy. It advocates and aspires to describe a socially responsible and scientifically respectable means of effecting social change, informed with empirical study and evaluation. It defines experimental social innovation and discusses in varying depth a wide sampling of problems and issues, from the planning of the administrative, scientific, and implementative components of that innovation, to its evaluation using both social and scientific criteria, and the training of the personnel therefor. A proposal for a center for research and training in experimental social innovation is advanced, and some of the professional, political, and social issues anticipated as a result of such a center are noted and discussed (e.g., a possible development, or fear of the development, of a behavioral technocracy as a political force). The book addresses the problem of generating and testing alternative solutions to social problems.

This is not a handbook and will not give the student the knowledge of experimental design and statistical analysis necessary for the competent execution of experimental social innovation. However, the conceptualization of the structure of this research is nicely presented, and the book brings the appropriate literature to the attention of the reader through liberal reference to more technical discussions. Some omissions are noted in this regard; the nearly standard reference to Campbell and Stanley's *Experimental and Quasi Experimental Designs for Research* (Chicago: Rand McNally, 1966) does not appear here, for example. Treatment of such other issues as "obtaining administrative commitments" may seem more impressive, unless that is what the reader is close to. Specialists in some particular phase of the program will find the treatment terse, perhaps sketchy, but will gain from reading sections further distant from their own concerns. The book addresses a wide domain, and therein lies its value. It brings between two covers the reflections of a scientist engaged in experimental social innovation, his outline of the choice points encountered, the alternatives, and his prescriptions.

Fairweather prescribes widely. For example, on preparing for the research he says that "researchers concerned about the problems of [Negroes] should spend as much time living in their situation as . . . in formulating theories about them" (p. 41); on forming the research team he says, "It is exceedingly important, since social innovative research is multidisciplinary and essentially empirical in nature, that those individuals who are committed to a single approach or theoretical position do not become involved in these experiments" (p. 65).

The book is important because social science will increasingly be asked to manage social innovation—or have it wrested from our hands. Fairweather is concerned that we do it, and do it well. His book should be widely read; it should also be widely amended as we make additions to a scientifically sound repertoire for dealing in social innovation. Since social experiments of the magnitude discussed will probably not easily be repeated, their planning and execution must be done right the first time. Social innovation is ongoing, and many will agree that it is accelerating. Fairweather points out that it is better implemented on the basis of knowledge of outcomes than in terms of politics. This book is a contribution toward the amassing of the requisite knowledge.

ROY S. MALPASS, *University of Illinois at Urbana-Champaign*

EDUCATIONAL PSYCHOLOGY: A COGNITIVE VIEW. By DAVID P. AUSUBEL. New York: Holt, Rinehart and Winston, 1968. Pp. xviii, 685. $8.95.

"If I had to reduce all of educational psychology to just one principle, I would say this: The most important single factor influencing learning is what the learner already knows. Ascertain this and teach him according" (p. vi). This, Ausubel's pre-introductory statement, is a challenging attempt to muster up sentiment, and at the same time it gives us a glimmer of the author's thinking. His text represents a significant contribution to the field. Evidence is presented with clarity, which should help the reader go beyond the statement just quoted; the book offers its reader substantive knowledge and yet leaves enough unsaid to generate question. Ausubel presents a well-documented and well-written account of a select number of important issues in educational psychology—an account which, one hopes, will help to uplift the level of instruction in educational psychology. Though relatively simply written, the text contains many complex ideas and appears aimed at the advanced student; because of this, it may not find its way into the hands of the great number of potential teachers who could benefit from its contents.

An overview of the text can be gleaned from the section headings: "Introduction," "Meaning and Meaningful Learning," "Cognitive Factors in Learning," "Affective and Social Factors in Learning," "Discovery Learning" (which includes excellent chapters on acquisition of concepts and on problem solving), and lastly, "Evaluation and Measurement." Hopefully a revision will strengthen this last section, which is of increased importance in light of Ausubel's initial statement.

To potential readers I suggest the scanning of two additional reviews of the text: W. R. Dixon's "Educational Psychology Texts: Two new models" (*Contemp. Psychol.*, 14, 1969, 421–423) and R. M. Gagne's (*Amer. educ. Res. J.* 6, 1969, 287–290). To any daring professor I suggest that a simple publishable experiment to test the efficacy of the "advanced organizers" Ausubel displays at the beginning of each chapter might help answer the query "Educational psychology: a cognitive view?"

HAROLD COOK, *Teachers College, Columbia University*

PARAPSYCHOLOGY TODAY. Edited by J. B. RHINE and ROBERT BRIER. New York: The Citadel Press, 1968. Pp. 286. $6.00.

It is likely that most psychologists who encounter this book will dismiss it as unimportant. If one considers only the empirical findings reported in this volume, there is considerable justification for such a response. All the problems that for decades have plagued psychical research are here in abundance: inadequate controls, obvious possibilities for experimenter bias, and unconvincing techniques for assessing psychic phenomena. These problems have been described in detail by others and need not be elaborated here.

Not only have these problems been discussed elsewhere, but at this particular point in history they may be irrelevant. As Cyril Burt notes in his contribution to the book, to accept at face value such phenomena as clairvoyance, precognition, telepathy, and so forth requires a reversal of several assumptions regarding the nature of causality and the materialistic foundation of mental processes. These assumptions comprise a major aspect of the world view held by most psychologists, and are shared by the other social and natural sciences. Although Burt discusses these assumptions in some detail, he does not carry the discussion to its inevitable conclusion, which is that to adopt the explanations of parapsychology brings one in conflict with several centuries of scientific accomplishment. In view of the enormous impact of these accomplishments on man's understanding and control of nature, it is unlikely that the assumptions on which they are based will be discarded on the strength of a handful of dubious counter-instances.

The real value of the book lies in two sections devoted to the views parapsychologists have of their discipline and of themselves. Included, for example, are chapters on the relationship between parapsychology and other domains of belief (religion and communism), as well as a chapter that reports the views of several "senior parapsychologists" on the fruitfulness of a career in parapsychology. It is in this section of the book that the truly important questions are touched upon—questions that social psychologists, for example, should be asking. That is, just who are the believers, what do they believe, and why? These are not idle questions. Parapsychology is flourishing, and it is not flourishing in isolation. Many diverse groups—ranging from hippie cults, to astrology devotees, to mystics of all persuasions—seem to share certain antimaterialistic beliefs with the parapsychologists. The overlap between the goals and beliefs of parapsychologists and others needs to be examined and understood, and it is from this perspective that these articles assume importance. The psychologist interested in studying social movements should not ignore parapsychology, nor should he ignore this book.

David A. Summers, *University of Kansas*

BOOKS RECEIVED

(The books listed here have not as yet been noted in our pages. Listing here does not, however, preclude their later review.)

Abramson, Harold A. *Psychological Problems in the Father-Son Relationship.* New York: October House, 1969. Pp. 223.

Andersson, Bengt-Erik. *Studies in Adolescent Behavior.* Stockholm: Almquist & Wiksell, 1969. Pp. 400.

Annett, John. *Feedback and Human Behavior.* Baltimore: Penguin Books, 1969. Pp. 196. $1.45.

Aubert, Vilhelm, ed. *Sociology of Law.* Baltimore: Penguin Books, 1969. Pp. 367. Paperback, $2.25.

Bales, Robert Freed. *Personality and Interpersonal Behavior.* New York: Holt, Rinehart and Winston, 1970. Pp. 561. $11.95.

Benton, Arthur L., ed. *Contributions to Clinical Neuropsychology.* Chicago: Aldine, 1969. Pp. 243. $8.50.

Berlin, Brent, and Kay, Paul. *Basic Color Terms: Their Universality and Evolution.* Berkeley and Los Angeles: University of California Press, 1970. Pp. 178. $8.50.

Beterille, Andre, ed. *Social Inequality.* Baltimore: Penguin Books, 1969. Pp. 414. Paperback, $2.25.

Birren, Faber. *Light, Color, and Environment.* New York: Van Nostrand, Reinhold, 1969. Pp. 131. $19.95.

Bradburn, Norman. *The Structure of Psychological Well-being.* Chicago: Aldine, 1970. Pp. 318. $9.75.

Bradford, Kirk A. *Existentialism and Casework.* New York: Exposition Press, 1969. Pp. 82. $4.50.

Brill, Leon, and Lieberman, Louis. *Authority and Addiction.* Boston: Little, Brown and Co., 1969. Pp. 318. $13.50.

Bullough, Bonnie. *Social-Psychological Barriers to Housing Desegregation.* Berkeley and Los Angeles: University of California Press, 1969. Pp. 133. $4.50.

Burgess, Robert L., and Bushell, Don, Jr. *Behavioral Sociology.* New York: Columbia University Press, 1969. Pp. 418. $12.50.

Burns, Tom, ed. *Industrial Man.* Baltimore: Penguin Books, 1969. Pp. 414. Paperback, $2.25.

Buros, Oscar Krisen, ed. *Personality Tests and Reviews.* Highland Park, N. J.: Gryphon Press, 1970. Pp. 1,690. $35.00.

Burris, Donald S., ed. *The Right to Treatment.* New York: Springer, 1969. Pp. 229. $3.75.

Cancro, Robert, ed. *The Schizophrenic Reactions.* New York: Brunner-Mazel, 1969. Pp. 293. $10.00.

Carkhuff, Robert R. *Helping and Human Relations: A Primer for Lay and Professional Helpers.* Vol. 1, *Selected Training.* New York: Holt, Rinehart and Winston, 1969. Pp. 298. $7.50.

Cattell, N. R. *The New English Grammar.* Cambridge, Mass.: The M.I.T. Press, 1969. Pp. 162. $8.95.

Cavitch, David. *D. H. Lawrence and the New World.* New York: Oxford University Press, 1969. Pp. 232. $6.00.

Chadwick-Jones, J. *Automation and Behaviour.* New York: Wiley, Interscience, 1970. Pp. 168. $8.95.

Chauncey, Henry, ed. *Soviet Preschool Education.* Vol. 2, *Teacher's Commentary.* New York: Holt, Rinehart and Winston, 1969. Pp. 218. Paperback, $4.95.

Chiang, Hung-Min, and Maslow, Abraham S., eds. *The Healthy Personality: Readings.* New York: Van Nostrand, Reinhold, 1969. Pp. 216. $2.75.

Chin, Robert, and Chin, Ai-li. *Psychological Research in Communist China, 1949–1966.* Cambridge, Mass.: The M.I.T. Press, 1969. Pp. 274. $10.00.

Clark, Alfred W., and Yeomans, Neville T. *Theory, Practice, and Evaluation of a Therapeutic Community: Fraser House.* New York: Springer, 1969. Pp. 282. $7.50.

Cole, Luella, and Hall, Irma Nelson. *Psychology of Adolescence.* 7th ed. New York: Holt, Rinehart and Winston, 1970. Pp. 669. $9.95.

Coles, Robert, and Piers, Maria. *Wages of Neglect.* Chicago: Quadrangle Books, 1969. Pp. 191. $5.95.

Costello, C. G., ed. *Symptoms of Psychopathology:* A Handbook. New York: Wiley, 1970. Pp. 679. $16.95.

Daniels, Roger, and Kitano, Harry. *American Racism: Exploration of the Nature of Prejudice.* Englewood Cliffs, N. J.: Prentice-Hall, 1970. Cloth, $4.95; paperback, $1.95.

Das, J. P. *Verbal Conditioning and Behavior.* Oxford: Pergamon Press, Ltd., 1970. Pp. 163. $7.00.

Delay, J., and Pichot, P. *Abrege de psychologie.* Paris: Masson et Cie, 1969. Pp. 490. 28 Fr.

Demos, John. *A Little Commonwealth.* New York: Oxford University Press, 1970. Pp. 201. $6.75.

Deutsch, Ronald M. *The Key to Feminine Response in Marriage.* New York: Random House, 1969. Pp. 172.

Dinsmoor, James A. *Operant Conditioning (19).* Dubuque, Iowa: Wm. C. Brown, 1969. Pp. 48. Paperback, $.48.

Duster, Troy. *The Legislation of Morality.* New York: Macmillan, 1970. Pp. 274. $6.95.

Eyres, A. E., and Pearson, C. T. *A Happier Life.* Durham, N. C.: Moore Publishing, 1969. Pp. 270. $6.95.

Foulkes, S. H., and Prince, Stewart G., eds. *Psychiatry in a Changing Society.* New York: Barnes & Noble, 1970. Pp. 211. $8.00.

Fowler, Orson Squire, and Fowler, Lorenzo Niles. *Phrenology: A Practical Guide to Your Head.* New York: Chelsea House, 1969. Pp. 201. $10.00.

Gilbert, Joseph. *Clinical Psychological Tests in Psychiatric and Medical Practice.* Springfield, Ill.: Charles C Thomas, 1969. Pp. 295. $10.00.

Goffman, Eriring. *Strategic Interaction*. Philadelphia: University of Pennsylvania Press, 1970. Pp. 145. $3.45.

Goldthorpe, John H.; Lockwood, David; Bechhofer, Frank; and Platt, Jennifer. *The Affluent Worker in the Class Structure*. Cambridge Studies in Sociology, No. 3. New York: Cambridge University Press, 1969. Pp. 239. Cloth, $7.50; paperback, $2.25.

Gordon, Ira J. *Human Development: From Birth through Adolescence*. 2d ed. New York: Harper & Row, 1969. Pp. 408. $8.95.

Gray, William; Buhl, Frederick J.; and Rizzo, Nicholas D.; eds. *General Systems Theory and Psychiatry*. Boston: Little, Brown and Co., 1969. Pp. 481. $15.00.

Gurr, Ted Robert. *Why Men Rebel*. Princeton, N. J.: Princeton University Press, 1970. Pp. 421. $12.50.

Halpern, Manfred. *Applying a New Theory of Human Relations to the Comparative Study of Racism*. Center of International Race Relations' Race and National Monograph Series, No. 1. Denver: University of Denver, 1969. Pp. 41.

Haymaker, Webb; Anderson, Evenlyn; and Nauta, Walle J. H.; eds. *The Hypothalamus*. Springfield, Ill.: Charles C Thomas, 1969. Pp. 805. $45.00.

Hazlitt, Henry. *Thinking as a Science*. Los Angeles: Nash Publishing, 1969. Pp. 162. $4.95.

Hill, John P., ed. *Minnesota Symposia on Child Psychology*. Vol. 3. Minneapolis: University of Minnesota Press, 1969. Pp. 163. $5.00.

Hofmann, Hans. *Discovering Freedom*. Boston: Beacon Press, 1969. Pp. 100. $6.00.

Hudson, Liam. *Frames of Mind*. New York: W. W. Norton, 1970. Pp. 134. $5.50.

Hyman, Herbert H. *Political Socialization: A Study in the Psychology of Political Behavior*. New York: Macmillan, 1970. Pp. 140. Paperback, $1.95.

Isard, Walter, and others. *General Theory: Social, Political, Economic, and Regional*. Cambridge, Mass.: The M.I.T. Press, 1969. Pp. 1,040. $10.00.

Jacobs, Jerry. *The Search for Help: A Study of the Retarded Child in the Community*. New York: Brunner-Mazel, 1969. Pp. 135. $5.95.

Jahoda, Gustav. *The Psychology of Superstition*. Baltimore: Penguin Books, 1969. Pp. 158. $6.95.

Jarvinen, Jaakko. *Orthogonal Localization of Visual Objects*. Helsinki: Academia Scientiarum Fennica, 1969. Pp. 141.

Jasper, Herbert H.; Ward, Arthur D.; and Pope, Alfred; eds. *Basic Mechanisms of the Epilepsies*. Boston: Little, Brown and Co., 1969. Pp. 864. $30.00.

Kahn, Herman. *On Thermonuclear War*. New York: Macmillan, 1969. Pp. 688. $3.95.

Kleinmuntz, Benjamin. *Clinical Information Processing by Computer: An Essay and Selected Readings*. New York: Holt, Rinehart and Winston, 1969. Pp. 399. $5.95.

Kolb, Lawrence C., and others, eds. *Urban Challenges to Psychiatry*. Boston: Little, Brown and Co., 1969. Pp. 512. $15.00.

Kroll, A. M., and others. *Career Development, Growth, and Crisis*. New York: Wiley, 1970. Pp. 262. $8.95.

Kronhausen, Phyllis, and Kronhausen, Eberhard. *Erotic Fantasies*. New York: Grove Press, 1970. Pp. 429. $10.00.

Langer, Jonas. *Theories of Development*. New York: Holt, Rinehart and Winston, 1970. Pp. 191. $5.95.

Lindesmith, Alfred R., and Anselm, Strauss L. *Readings in Social Psychology*. New York: Holt, Rinehart and Winston, 1969. Pp. 357. Paperback, $6.95.

Lorenz, Konrad. *Studies in Animal and Human Behaviour*. Vol. 1. Cambridge, Mass.: Harvard University Press, 1970. Pp. 403. $10.00.

Lowe, Gordon R. *Personal Relationships in Psychological Disorders*. Baltimore: Penguin Books, 1969. Pp. 143. $1.25.

McDill, Edward L.; McDill, Mary S.; and Sprabe, J. Timothy. *Stragtegies for Success in Compensatory Education*. Baltimore: The Johns Hopkins Press, 1970. Pp. 83. Cloth, $5.00; paperback, $1.95.

McElhaney, Mark. *Clinical Psychological Assessment of the Human Figure Drawing*. Springfield, Ill.: Charles C Thomas, 1969. Pp. 256. $13.00.

MacKay, Donald M. *Information, Mechanism and Meaning*. Cambridge, Mass.: The M.I.T. Press, 1969. Pp. 196. Cloth, $8.95; paperback, $2.95.

Miller, George A. *The Psychology of Communication*. Baltimore: Penguin Books, 1969. Pp. 197. $1.25.

Morris, Desmond. *The Human Zoo*. New York: McGraw-Hill, 1969. Pp. 256. $6.95.

Myers, P. I., and Hammill, D. D. *Methods for Learning Disorders*. New York: Wiley, 1969. Pp. 313. $8.50.

Naumenko, A. I., and Benua, N. N. *The Physiological Mechanisms of Cerebral Blood Circulation*. Springfield, Ill.: Charles C Thomas, 1970. Pp. 123. $8.75.

Neubec, Gerhard, ed. *Extramarital Relations*. Englewood Cliffs, N. J.: Prentice-Hall, 1970. Pp. 205. Cloth, $5.95; paperback, $2.45.

Neumann, Erich. *Depth Psychology and a New Ethic*. New York: G. P. Putnam's Sons, 1970. Pp. 158. $7.00.

O'Donnell, John A. *Narcotic Addicts in Kentucky*. Chevy Chase, Md.: National Clearinghouse for Mental Health Information, 1969. Pp. 297.

Payne, Dorothy. *Women without Men*. Philadelphia: Pilgrim Press, 1969. Pp. 150. $4.95.

Perry, Nathan W., and Childers, Donald G. *The Human Visual Evoked Response: Method and Theory*. Springfield, Ill.: Charles C Thomas, 1969. Pp. 187. $14.00.

Pribram, Karl, ed. *Adaptation*. Baltimore: Penguin Books, 1969. Pp. 524. Paperback, $2.95.

Pribram, Karl, ed. *Brain and Behavior*. Vol. 1, *Mood, States, and Mind*. Vol. 2, *Perception and Action*. Vol. 3, *Memory Mechanisms*. Baltimore: Penguin Books, 1969. Pp. 496, 575, 524. Paperback, $2.95 each.

Rabkin, Richard. *Inner and Outer Space*. New York: W. W. Norton, 1970. Pp. 215. $7.50.

Reed, Homer B. *History of Psychological Associations in Kansas*. Hays: Kansas Psychological Association, 1969. Pp. 36.

Robertson, Roland, ed. *Sociology of Religion*. Baltimore: Penguin Books, 1969. Pp. 473. $2.25.

Rose, Richard. *Policy Making in Britain*. New York: The Free Press, 1969. Pp. 375. $9.95.

Roslansky, John D., ed. *Communications: Nobel Conference Lectures*. New York: Fleet Academic Editions, 1970. Pp. 131. $6.95.

Rubin, Melvin L., and Walls, Gordon L. *Fundamentals of Visual Science.* Springfield, Ill.: Charles C Thomas, 1969. Pp. 435. $17.50.

Ruitenbeek, Hendrik M., ed. *Group Therapy Today.* New York: Atherton Press, 1969. Pp. 365.

Ruwet, Jean Claude. *Ethologie: Biologie du comportement.* Brussels: Charles Dessart, 1969. Pp. 238.

Rycroft, Charles. *Anxiety and Neurosis.* Baltimore: Penguin Books, 1968. Pp. 160. $5.95.

Salzinger, Kurt. *Psychology, the Science of Behavior.* New York: Springer, 1969. Pp. 306. $3.95.

Schmidhofer, Ernst. *Cerebral Training.* New York: Hafner, 1969. Pp. 178. $12.50.

Shepard, Marin, and Lee, Marjorie. *Games Analysts Play.* New York: G. P. Putnam's Sons, 1970. Pp. 189. $5.95.

Shulman, Lee, and Taylor, Joan Kennedy. *When to See a Psychologist.* Los Angeles: Nash Publishing, 1970. Pp. 311. $6.95.

Siegel, A. I., and Wolf, J. J. *Man-Machine Simulation Models.* New York: Wiley, Interscience, 1969. Pp. 177. $9.95.

Spetz, Marseille. *The Tale of the Seventh Trumpet.* Los Angeles: Deuce of Clubs Press, 1969. Pp. 64. $5.00.

Stewart, J. H. *Premarital Counseling for Teen Agers.* New York: Exposition Press, 1969. Pp. 45. $3.00.

Stone, Le Roy A. *Psychophysical Scaling Applications with Clinical and Educational Concerns.* New York: MSS Educational Publishing, 1969. Pp. 104. $3.75.

Sturdivant, Frederick D., ed. *The Ghetto Marketplace.* New York: The Free Press, 1969. Pp. 316. $7.95.

Suinn, Richard M., and Oskamp, Stuart. *The Predictive Validity of Projective Measures: A Fifteen-Year Evaluative Review of Research.* Springfield, Ill.: Charles C Thomas, 1969. $9.50.

Turner, Dean E. *The Autonomous Man.* Saint Louis: Bethany Press, 1970. Pp. 206. $4.95.

Walker, Helen M., and Lev, Joseph. *Elementary Statistical Methods.* 3d ed. New York: Holt, Rinehart and Winston, 1969. Pp. 432. $8.95.

Ward, Charles D. *Laboratory Manual in Experimental Psychology.* New York: Holt, Rinehart and Winston, 1970. Pp. 468. $6.50.

Weinstein, Fred, and Platt, Gerald M. *The Wish to Be Free: Society, Psyche, and Value Change.* Berkeley and Los Angeles: University of California Press, 1970. Pp. 319. $8.50.

Willmott, Peter. *Adolescent Boys of East London.* Rev. ed. Baltimore: Penguin Books, 1969. Pp. 237. $1.45.

Wolf, James M., ed. *The Results of Treatment in Cerebral Palsy.* Springfield, Ill.: Charles C Thomas, 1969. Pp. 327. $20.50.

Wolff, Sula. *Children under Stress.* Baltimore: Penguin Books, 1969. Pp. 249. $6.95.

Wrage, Karl. *Man and Woman.* Philadelphia: Fortress Press, 1969. Pp. 259. $8.75.

Yates, Aubrey J. *Behavior Therapy.* New York: Wiley, 1970. Pp. 445. $13.95.

Diary N 189
Date 8 . 1 . 71
File lib ~
Edn'l. Psy. Research

A TEST OF THE LIST-DIFFERENTIATION HYPOTHESIS

EDWIN MARTIN and STEPHEN A. MACKAY,
University of Michigan, Ann Arbor

Abstract. The idea that original and interpolated paired-associate learning in a negative transfer paradigm produce list-differentiated response sets was tested by cluster analyzing unstimulated, freely recalled responses. List-determined recall organization failed to occur both in immediate and in delayed recall under the S_1–R_1, S_1–R_2 interference paradigm but did occur under the S_1–R_1, S_2–R_2 control paradigm.

The essence of the list-differentiation hypothesis is that in a paired-associate situation, original S_1–R_1 learning produces a permissible response set $\{R_1\}$, that interpolated S_1–R_2 learning produces $\{R_2\}$, and that the exercise of $\{R_2\}$ produces a suppression of $\{R_1\}$ which is assumed to underlie the retroactive-interference phenomenon. As time elapses from the end of interpolated S_1–R_2 learning, the suppression of $\{R_1\}$ dissipates to the point that R_1 and R_2 can both be recalled without mutual interference upon presentation of S_1. This dissipation of $\{R_1\}$ suppression is posited to produce the sometimes-observed phenomenon of spontaneous recovery.

The origin of the list-differentiation hypothesis is usually given to Thune and Underwood. They discussed the possibility of covert intrusion of R_2s in a subsequent test for what was learned in the S_1–R_1 task as partly responsible for retroactive interference, as "suggested by [subjects] who report, 'I thought of a response but I knew it was wrong, though I didn't have time to think of the correct response.' "[1] Undoubtedly this is what led Underwood to postulate that "the amount and character of this interference is . . . a function of two hypothetical constructs or intervening variables:

Received for publication March 6, 1970. This research was supported by the Advanced Research Projects Agency, Department of Defense, and monitored by the Air Force Office of Scientific Research, under Contract AF-49(638)-1735 with the Human Performance Center, Department of Psychology.

[1] L. E. Thune and B. J. Underwood, Retroactive inhibition as a function of degree of interpolated learning, *J. exp. Psychol.*, 32, 1943, 185–200.

[*a*] the associative strength of the IL [interpolated list] responses and [*b*] the degree of differentiation of the competing response systems."[2] Our immediate concern is with the second of these two factors. Differentiation was seen by Underwood as " 'knowing' the appropriateness of a 'thought-of response' " (p. 25) and was assumed to vary inversely with delay of the test for what was learned in the original S_1–R_1 learning. The idea is that at the end of interpolated S_1–R_2 learning the subjects 'know' that the R_2s are appropriate, and because the R_1s are equally known to be inappropriate, the R_1s are not given; hence the retroactive-interference phenomenon.

In 1960, Underwood and Schulz asked why overt R_1 intrusions during S_1–R_2 learning were as infrequent as they are. They then restated Underwood's 'appropriateness' notion by concluding that "the evidence is overwhelming that a selective response mechanism exists."[3] They wrote: "The selector mechanism, itself unexplained, is assumed here as a fundamental behavioral law. . . . In the verbal-learning situation, the selector mechanism restricts the spew [emission] to those response units within the list" (p. 144). Thus original S_1–R_1 learning produces the set $\{R_1\}$, and interpolated S_1–R_2 learning produces $\{R_2\}$.

That $\{R_1\}$ is suppressed or inhibited as the result of forming the set $\{R_2\}$ was first posited by Underwood in 1945 in terms of the subject's presumed knowledge of inappropriateness of the R_1s, and restated in 1968 by Postman: "RI [retroactive interference] on an unpaced test of recall is largely a matter of reduced response availability . . . specific associations are in general highly resistant to unlearning."[4] Postman elaborated thus: "(*a*) unlearning results from the operation of a mechanism of response selection under conditions of negative transfer and (*b*) this mechanism exerts its primary effect on the entire class of first-list responses rather than on specific stimulus-response associations" (p. 689). With regard to spontaneous recovery over time, Postman asserted that "the inertia which characterizes the selector mechanism is assumed to be of limited duration, that is, the degree of dominance of the most recent criteria of selection diminishes over time" (p. 691). We may note here that Postman acknowledged as background the response-selector mechanism of Underwood and Schulz, and the notion of generalized response competition as proposed by Newton and Wickens.[5] The latter is the idea that in transferring from one task to another, the subjects will persist for some time in their subjective set to emit those responses they have just been emitting.

---

[2] B. J. Underwood, The effect of successive interpolations on retroactive and proactive inhibition, *Psychol. Monogr.*, 59, 1945, No. 3 (Whole No. 273), 24.

[3] B. J. Underwood and R. W. Schulz, *Meaningfulness and Verbal Behavior*, 1960, 143.

[4] L. Postman, K. Stark, and J. Fraser, Temporal changes in interference, *J. verb. Learn. verb. Behav.*, 7, 1968, 689.

Thus according to the most recent version of the list-differentiation hypothesis, original S_1–R_1 learning produces the set $\{R_1\}$ by action of a response-selector mechanism; interpolated S_1–R_2 learning likewise produces the set $\{R_2\}$. Negative transfer is thus ascribable to list-differentiation learning, that is, to the process of differentiating $\{R_1\}$ and $\{R_2\}$. This is implied by the assumption that the subjective set for $\{R_1\}$ has an inertial characteristic (see nn. 5, 4). That interpolated learning is assumed not to interfere with original S_1–R_1 associations (see n. 4), together with the assumption that original and interpolated learning produce the permissible response sets $\{R_1\}$ and $\{R_2\}$, means that at the end of interpolated learning there should be in memory a list-membership organization of all the responses experienced in the two successive tasks. This inference is independently supported by numerous reports of subjects' ability to accurately identify the learning task from which a given response comes.[6]

It is well known that memorial organization as imposed by a variety of experimental and normative procedures has a strong and marked effect on the organization of output in recall.[7] We have argued that a direct implication of the list-differentiation hypothesis is that responses are organized in memory according to list membership. From these two considerations in conjunction, it follows that a cluster analysis of free, unstimulated recall of the responses of both lists should reveal above-chance list clustering. List organization of recall should be sharply evident in immediate recall, but evident only weakly or not at all in delayed recall.

---

[5] J. M. Newton and D. D. Wickens, Retroactive inhibition as a function of the temporal position of interpolated learning, *J. exp. Psychol.*, 51, 1956 149–154.

[6] E.g., I. M. Birnbaum, Long-term retention of first-list associations in the A–B, A–C paradigm, *J. verb. Learn. verb. Behav.*, 4, 1965, 515–520; R. J. Koppenaal, Time changes in the strengths of A–B, A–C lists, *J. verb. Learn. verb. Behav.*, 2, 1963, 310–319.

[7] E.g., W. A. Bousfield, The occurrence of clustering in the recall of randomly arranged associates, *J. gen. Psychol.*, 49, 1953, 229–240; B. H. Cohen, Recall of categorized word lists, *J. exp. Psychol.*, 66, 1963, 227–234; G. Mandler, Organization and memory, in K. W. Spence and J. T. Spence (eds.), *The Psychology of Learning and Motivation*, 1967, 327–372.

It makes sense to include the S_1–R_1, S_2–R_2 control paradigm in a study where the responses are to be freely recalled after interpolated learning. In the S_1–R_1, S_2–R_2 paradigm, there is no basis for the alternative—clustering on the basis of common stimuli—to list-determined clustering. This means that the clustering measure of recall organization can be compared with other measures of retroaction as a function of the presence or absence of common stimuli.

METHOD

Each *S* learned to criterion two successive lists of nine letter-word pairs by the anticipation method. For half of the *S*s, the two lists followed the S_1–R_1, S_1–R_2 paradigm; for the other half, the S_1–R_1, S_2–R_2 paradigm. Then, either immediately or 30 min. later, they were asked to (free) recall, verbally, all of the responses (words). The stimuli were not present during response recall. Thus the design was a 2 × 2, completely between-subjects factorial for two transfer paradigms and two response-recall intervals.

Subjects. Forty-eight women, students at the University of Michigan, were randomly assigned to the four conditions, with the single restriction that the number of *S*s in one condition never exceed the number in any other condition by more than two at any time. The *S*s volunteered for paid participation.

Materials. Two nonoverlapping sets of nine letters were selected so that the members of each set spanned the alphabet and that no two letters in either set formed an adjacent alphabetic sequence. These letters were the *stimuli* and were variously paired with sets of common words that served as responses. The lists of *responses* were drawn from two different pools, each consisting of 18 unrelated words (nouns, verbs). Each pool was partitioned into two sets, R_1 and R_2, in two different ways, yielding four different R_1, R_2 combinations. As with the stimulus sets, the first letters of the words in each response list came from nonadjacent and scattered positions of the alphabet. Four different S_1–R_1, S_1–R_2 and four different S_1–R_1, S_2–R_2 sequences of stimulus-response pairing assignments were constructed. Likely associations and other mnemonic possibilities between the stimuli (letters) and the responses (words) were carefully avoided.

Procedure. Before both the original and the interpolated tasks, the *S*s were given typical anticipation-method instructions to learn as quickly as possible to a criterion of once through correct. The *S*s were also told the nature of the interpolated task before beginning it. The learning materials were presented on a Stowe memory drum at a 2:2-sec. rate with a 2-sec. intertrial interval. Following interpolated learning, the *S*s whose recall was to be delayed for 30 min. were given a paper-and-pencil task in which they were to partition sentences into word groups according to self-selected subjective criteria of phrasing. At the time of recall, which followed either immediately after interpolated learning or after the 30-min. paper-and-pencil task, the *Ss* were instructed to recall all the words (responses) learned in the two paired-associate tasks. The instructions emphasized that the words should be recalled in whatever order they came to mind. Recall was verbal, and the *Ss* were told not to worry about repeating previously recalled words. There was no time limit on recall. After *S* completed his recall, he was asked to recall the words a second time. At no time in either recall attempt were the stimuli present.

RESULTS

Learning and transfer. Given both that two trials was the number required by a few subjects to reach criterion on interpolated learning and that specific transfer effects are adequately characterizable by early-trials performance in the interpolated task, the learning and transfer results are represented by the number of errors committed (out of 18 possible per subject) on the first two anticipation trials of the two tasks.

The mean number of Trial 1–2 errors in original learning was 11.0 in the S_1–R_1, S_1–R_2 paradigm; it was 12.2 in the S_1–R_1, S_2–R_2 paradigm. The mean number of Trial 1–2 errors in interpolated learning was 11.5 and 9.3 for these two paradigms respectively. Each subject's percent transfer from original to interpolated learning was calculated. The mean transfer score in the S_1–R_1, S_1–R_2 paradigm was −1.2; in the S_1–R_1, S_2–R_2 paradigm, 12.1. There was, thus, negative transfer in the former paradigm relative to the latter paradigm [$t\ (46) = 2.61, p < .02$].

Recall frequencies. The mean number of responses recalled from the two lists, original and interpolated, at each of the recall intervals for the two transfer paradigms, is shown in Table I. In this representation, those responses that were recalled twice were counted only once.

Analysis of variance of the first-recall data yielded a significant delay effect [$F\ (1,\ 44) = 31.51,\ p < .001$], a significant list (original, interpolated) effect [$F\ (1, 44) = 48.21, p < .001$], and a significant list-by-delay interaction [$F\ (1, 44) = 11.22, p < .005$]. The same analysis of the second-recall data yielded corresponding results [$F\ (1, 44) = 4.86, p < .05$; $F\ (1, 44) = 35.94, p < .001$; and $F\ (1, 44) = 3.99, .05 < p < .10$]. All remaining *F*s were less than one. These results mean that the original R_1s were less avail-

TABLE I

MEAN NUMBER OF RESPONSES RECALLED

	Original (R_1)		Interpolated (R_2)	
Paradigm	Immed.	30 min.	Immed.	30 min.
First recall				
S_1–R_1, S_1–R_2	5.5	7.4	8.0	8.3
S_1–R_1, S_2–R_2	5.2	7.2	7.8	8.2
Second recall				
S_1–R_1, S_1–R_2	5.9	7.2	7.7	8.1
S_1–R_1, S_2–R_2	5.4	6.8	7.5	7.8

able for recall in the immediate test than were the interpolated R_2s, and that the R_1s recovered in availability over 30 min. Moreover, this recall pattern was the same for the two paradigms.

The sentence-partitioning task that filled the 30-min. delay interval pertained to another research project. In a post factum examination of these sentences for words in common with words used as responses in original learning, we discovered five such words. Either none, one, or two of them appeared in List 1 according to the subject's list assignment. They accounted for 19% of all possible recalls in each of the four conditions. In actual recall, they accounted for 18% of the 30-min. first recalls both in the S_1–R_1, S_1–R_2 paradigm and in S_1–R_1, S_2–R_2 paradigm. The corresponding observed percentages in immediate first recall, which was not preceded by sentence partitioning, were 21% and 16%. Thus sentence partitioning was not in itself responsible for spontaneous recovery of the original R_1s.

Recall order. After deleting the second occurrence of those responses recalled twice, each response recalled was assigned a rank order according to its output position. Then the mean output position for the original R_1 responses was figured, as was the mean output position for the interpolated R_2 responses. The latter was then subtracted from the former, thus giving, for each subject, a mean difference in recall position for original and interpolated responses. The means of these differences are posted in Table II, where the value 3.1, for example, means that on the average interpolated responses preceded original responses by 3.1 serial positions in recall.

TABLE II

MEAN RANK-ORDER DIFFERENCE BETWEEN ORIGINAL AND INTERPOLATED RESPONSES IN RECALL

Paradigm	Immed.	30 min.
First recall		
S_1–R_1, S_1–R_2	3.1	−.2
S_1–R_1, S_2–R_2	3.7	1.5
Second recall		
S_1–R_1, S_1–R_2	.3	−1.1
S_1–R_1, S_2–R_2	−1.4	−.3

Analysis of variance of the first-recall data indicated that the only significant effect on recall order was due to the delay interval [$F(1, 44) = 7.18$, $p < .02$]. Transfer paradigm did not affect recall order. The same analysis of the second-recall data yielded no effects traceable either to delay or to transfer paradigm. Thus only in immediate first recall were original and interpolated responses different in recall order, with interpolated responses preceding original responses by approximately 3.4 serial positions on the average. This difference in output order in immediate first recall disappeared completely in the second recall.

Recall clustering. If a given subject recalls m_1 original responses and m_2 interpolated responses, then it follows from combinatorial analysis[8] that the expected number of adjacencies of recalls from the same list is given by

$$E(R) = \frac{m_1^2 + m_2^2}{m_1 + m_2} - 1.$$

Using this formula, we calculated $E(R)$ for each subject. Observed R was then determined for each subject by counting the actual number of times two responses from the same list were adjacent. In making this analysis, responses that were repeated in recall were retained. To include repeated recalls makes sense in that we are interested in the *organization* of recall, not the isolated occurrence or reoccurrence of individual responses.

If we let $n = m_1 + m_2$, and let

$$F_2 = \sum_{i=1}^{2} m_i(m_i - 1)$$

---

[8] E.g., F. N. David and D. E. Barton, *Combinatorial Chance*, 1962, Ch. 8.

and

$$F_3 = \sum_{i=1}^{2} m_i(m_i - 1)(m_i - 2),$$

then

$$\mathrm{Var}\,(R) = \frac{F_2(n-3)}{n(n-1)} + \frac{F_2^2}{n^2(n-1)} + \frac{2F_3}{n(n-1)}.$$

Thus for each subject we can calculate

$$r = \frac{R - E(R)}{\sqrt{\mathrm{Var}\,(R)}},$$

a sort of individual standardized score where $E(r)$ is zero. Note that r can be negative if a subject clusters less than expected.

In Table III, on the left, is shown the mean r for list-membership clustering for each of the various experimental conditions. For immediate first recall in the experimental S_1–R_1, S_1–R_2 paradigm, mean r was .50, which does not differ from zero [$t\ (11) = .82, p > .40$]. This means that in immediate recall there was no clustering evidence for list organization of the R_1 and R_2 responses. The difference $(.50 - .16 = .34)$ over the 30-min. delay was also not significant [$t\ (22) = .42,\ p > .60$]. In the S_1–R_1, S_2–R_2 control paradigm, however, immediate first recall yielded significant list-determined clustering [$t\ (11) = 3.74,\ p < .01$], but again the difference $(1.33 - 1.16 = .17)$ over the 30-min. delay was not significant [$t\ (22) = .31,\ p > .70$].

Analysis of variance of the individual r values that contributed

TABLE III

STANDARDIZED SCORE (r) FOR LIST AND STIMULUS CLUSTERING IN RECALL

Paradigm	List r Immed.	List r 30 min.	Stimulus r Immed.	Stimulus r 30 min.
First recall				
S_1–R_1, S_1–R_2	.50	.16	1.13	2.92
S_1–R_1, S_2–R_2	1.16	1.33		
Second recall				
S_1–R_1, S_1–R_2	.56	−.06	2.14	3.97
S_1–R_1, S_2–R_2	.44	.37		

to the four means in the upper left quadrant of Table III indicated a marginal inference for a paradigm effect [$F\ (1, 44) = 3.50$, $.05 < p < .10$]. For the delay-of-recall variable and its interaction with the paradigm factor, both Fs were less than one. With regard to second recall (lower left quadrant of Table III), the list r means differed neither from zero nor among themselves.

The recall protocols of the S_1–R_1, S_1–R_2 paradigm can be analyzed a second time using common stimuli as the criterion for clustering. In this analysis, we have

$$E(R) = \frac{m_1^2 + m_2^2 + \cdots + m_k^2}{n} - 1,$$

where k is the number of different stimuli (S_1s) for which either the original or the interpolated response was recalled, m_i is either one or two according to whether one or both responses were recalled, and n is the total number of responses recalled. Thus each subject has his own $E(R)$, and it is different from his $E(R)$ in the list-determined cluster analysis. The values for Var (R) also differ; in this analysis the summations run to k (instead of 2), F_2 is $2x$ where x is the number of stimuli for which both responses were recalled, and F_3 is zero. As before, a standardized score, r, was calculated for each subject.

The resulting mean r values are shown on the right in Table III. The 1.13 mean r value for immediate first recall in the S_1–R_1, S_1–R_2 paradigm was not reliably different from zero [$t\ (11) = 1.56$, $p > .10$]. However, for 30-min. first recall, stimulus-determined clustering was different from zero [$t\ (11) = 3.01, p < .02$]. For both immediate and 30-min. recall, the second-recall mean r values differed from zero [$t\ (11) = 3.96$ and 3.31 respectively, $p < .01$].

DISCUSSION

The principal result is that the succession of original and interpolated learning did not produce list-determined organization of unstimulated response recall in the S_1–R_1, S_1–R_2 interference paradigm. However, before concluding that this result denies the list-differentiation hypothesis in its present form, the prior credentials for this result must be stated. First of all, transfer in the interference paradigm was negative relative to the control paradigm. Second, retroactive interference was observed. It is well known that in a rest control condition where the interpolated task is omitted, general forgetting of the original responses is minuscule;[9] yet in the present experiment, only 5.5 original responses out of 9 possible were recalled in the first immediate test. Third, there was absolute spontaneous recovery. In the first recall, original response output increased from 5.5 to 7.4 out of 9 over the 30-min. recall interval. Fourth, in the first immediate recall, the interpolated responses were highly available, 8.0 out of 9, and preceded the original responses in recall by 3.1 serial positions on the average. But in the first 30-min. recall, the interpolated responses were only slightly more available than the original responses and did not precede them in output. These facts clearly indicate that the original R_1s suffered substantial retroactive interference and that they recovered from this interference during the 30-min. delay. Nevertheless, the amount of list-determined clustering was not distinguishable from that expected by chance alone at either recall interval.

That significant list clustering occurred under the S_1–R_1, S_2–R_2 paradigm argues that the test procedure itself was not responsible for the absence of list clustering under the S_1–R_1, S_1–R_2 paradigm. A further datum against the current form of the list-differentiation hypothesis is the fact that even though in the S_1–R_1, S_2–R_2 paradigm there was R_1 unavailability in immediate recall[10] and recovery from unavailability over 30 min., there was no corresponding change in the amount of list-determined clustering.

The logical alternative to list clustering in the S_1–R_1, S_1–R_2 paradigm is clustering on the basis of common stimuli. We are not in a position to elaborate on this possibility because stimulus clustering did not differ significantly from chance in the first immediate recall (1.13 in Table III), and although stimulus clustering in the first 30-min. recall (2.92 in Table III) differed from chance, it did not differ significantly from that in the first immediate recall. Moreover, while stimulus clustering exceeded list clustering in second recall [F (1, 22) $= 8.03$, $p < .01$], it did not in first recall [F (1, 22) $= 1.61$, $p > .20$]. If any conclusion is to be drawn from this pattern of results, it might be that stimulus organization emerges upon the subject's being given a second chance to recall the responses.

---

[9] E.g., Postman, Stark, and Fraser (see n. 4 above); G. Keppel, D. M. Henschel, and B. Zavortink, Influence of nonspecific interference on response recall, *J. exp. Psychol.*, 81, 1969, 246–255.

[10] Keppel, Henschel, and Zavortink (n. 9), 249–253, have previously established that when the test for retroaction is free recall of the responses, there is as much R_1 unavailability in the S_1–R_1, S_2–R_2 paradigm as in the S_1–R_1, S_1–R_2 paradigm.

It is certainly the case that a subject who did organize his recall primarily on the basis of common stimuli would then be limited on how much stimulus clustering he could produce, and vice versa. Indeed, over the 12 subjects in the immediate-recall condition, the product-moment correlation between stimulus- and list-clustering scores in first recall was $-.75$ ($p < .01$). The corresponding correlation for the 30-min. first recall was $-.87$ ($p < .01$). This means there was a close trade-off between the two types of recall organization, which, together with the fact that the average subject does not list-cluster his responses in recall, implies that the list-differentiation hypothesis only occasionally describes the behavior of a particular subject, and then an atypical one.

It is, of course, possible to counter this conclusion by saying that list clustering in free recall is not something the formulators of the list-differentiation hypothesis had in mind. However, it is nevertheless a valid implication of the hypothesis; hence denial of the implication should serve to sharpen the specifics of the hypothesis. Here is one possible modification. Responses are organized into sets according to the most salient feature of the experimental situation. In the S_1–R_1, S_2–R_2 control paradigm, the most salient feature is two distinct lists; hence list clustering occurs in first recall. But given that first recall has occurred, the task has changed so that in second recall list differentiation is no longer a salient feature. In the S_1–R_1, S_1–R_2 interference paradigm, there are two salient features: list membership and common stimuli. At no time, however, is list membership the dominant feature for the typical subject; indeed, in 30-min. recall and in second recall at either delay, stimulus organization seems to dominate.

ADAPTATION TO UNIOCULAR IMAGE MAGNIFICATION: MODIFICATION OF THE DISPARITY-DEPTH RELATIONSHIP

WILLIAM EPSTEIN and CYNTHIA LAVERNE MORGAN,
University of Wisconsin, Madison

Abstract. Experimental subjects were exposed to a meridional size lens for one hour. Measures of adaptation and aftereffect showed significant adaptive shifts. Control subjects, who were exposed to nondistorting plain glass for one hour, did not show shifts. The results are interpreted as a revaluation of the relationship between binocular disparity and perceived depth as a result of exposure to informational discrepancy.

Experiments by Wallach et al. and Epstein have shown that durative exposure to informational discrepancy, involving binocular disparity and the kinetic depth effect, leads to modification of the disparity-depth relationship.[1] The present experiment investigated the modifiability of the disparity-depth relationship with a new procedure.

A meridional size lens (MSL) before one eye magnifies the image in that eye along the meridian perpendicular to the axis of the lens. With the lens's axis at 90° (vertical), the image is magnified in the horizontal meridian. Uniocular horizontal magnification changes the disparities between the images of the two eyes, causing highly predictable perceptual distortions. Ames has provided verbal and graphic descriptions of the perceptual effects.[2] One of these effects was measured in the present experiment. The frontal plane

---

Received for publication December 29, 1969. This investigation was supported by research grant MH-16399-02 from the National Institute of Mental Health.

[1] H. Wallach, M. E. Moore, and L. Davidson, Modification of stereoscopic depth perception, this JOURNAL, 76, 1963, 191–204; W. Epstein, Modification of the disparity-depth relationship as a result of exposure to conflicting cues, this JOURNAL, 81, 1968, 189–197.

[2] A. Ames, Jr., Binocular vision as affected by relations between uniocular stimulus-patterns in commonplace environments, this JOURNAL, 59, 1946, 333–357.

appears to be rotated about a vertical axis through the fixation point. A surface in the objective frontal plane appears slanted, the side in front of the eye with the lens seeming farther away. This effect is well understood, and Ogle has derived an equation to predict the magnitude of apparent rotation when conflicting depth cues are absent.[3] A number of investigators have reported good agreement between apparent slant and the predictions provided by Ogle's equations.[4]

When a subject views a typical environmental setting with an MSL before one eye, he is exposed to informational discrepancy concerning the spatial arrangement of the environment. For example, disparity information indicates that the front wall is slanted back, while information provided by monocular depth cues plus everyday experience suggests that the wall is frontal-parallel. If a subject wears the lens continuously, will the informational discrepancy lead to a reevaluation of the disparity cue? An early study by Burian suggested that reevaluation may occur.[5] Our experiment was designed to study this question in an improved version of the contemporary adaptation experiment.[6]

METHOD

Subjects. The *S*s were 42 undergraduates with uncorrected vision. They were paid for serving in the experiment.

Apparatus. Two pairs of melter's *goggles* were used. The pair worn by the experimental *S*s had a 5% MSL, axis 90°, in one frame and plain glass in the other frame. The pair worn by the control *S*s had plain glass in both frames. The *depth-discrimination apparatus* used is described in detail by Ebenholtz and Walchli.[7] Briefly, two fine electroluminescent vertical lines, separated laterally by 65 mm., are presented in total darkness. By turning a knob, *E* causes the two lines to move slowly and continuously, one advancing and the other receding by exactly the same amount. The lines are visible to *S* through a rectangular aperture at the face of the apparatus. The dimensions of the aperture ensure that the visible portions of both lines subtend the same, invariant visual angle at all positions of the lines in the apparatus.

---

[3] K. Ogle, *Researches in Binocular Vision*, 1964.

[4] Ogle (see n. 3 above); B. Gilliam, Perception of slant when perspective and stereopsis conflict, *J. exp. Psychol.*, 78, 1968, 299–305.

[5] M. Burian. Influence of prolonged wearing of meridional size lenses on spatial localization, *Arch. Opthalmol.*, 30, 1943, 645–666.

[6] W. Epstein, *Varieties of Perceptual Learning*, 1967, Ch. 9.

[7] S. M. Ebenholtz and R. M. Walchli, Stereoscopic thresholds as a function of head and object orientation, *Vis. Res.*, 5, 1965, 455–456 ff.

When *S*'s head is stationary and the room is dark, the discrimination of depth intervals between the lines depends exclusively on binocular disparity. Informal tests restricting the observer to monocular vision supported this claim by showing that depth discrimination was not possible without binocular vision.

Procedure. The typical design of the adaptation experiment was elaborated in three ways. First, direct measures of adaptation were secured, as well as measures of the aftereffect of adaptation. This was accomplished by obtaining measures of change with MSL on, as well as with MSL absent. Second, procedures were employed that prevented *S* from knowing whether he was looking through the distorting goggles (MSL) or the nondistorting goggles (plain glass) during the test. Third, a control condition was introduced, in which *S*s followed the same procedure as experimental *S*s, except that these control *S*s wore plain glass goggles during the adaptation period.

The *S* was seated with his head in a restraining device that prevented head movement. The *S*'s head was frontal-parallel with the viewing aperture, his eye-level at the height of the midpoint of the aperture. The distance from the eyes to the equidistance position of the two test lines was 6 ft. The *S*s were assigned to experimental and control conditions in regular alternation. In the first stage of the procedure, three sets of four equidistance settings were secured using the method of adjustment. First, four judgments were obtained with plain glass goggles. Two of the trials began with the right line more distant, and two trials with the left line more distant. Then eight judgments were obtained with the MSL, four with the lens in front of the right eye and four with the lens in front of the left eye. This procedure required two changes of goggles. The changes were accomplished while *S*'s eyes were closed and without any indication by *E* about the nature of the glasses. After the last of the 12 preadaptation judgments, the MSL goggles were removed again. They were replaced by plain glass goggles for the control *S*s. For the experimental *S*s the MSL goggles were used again, with the MSL in front of the right eye for half of the *S*s and in front of the left eye for the other *S*s.

A one-hour exposure period followed. During this period all *S*s walked the corridors and stairs of the Psychology Building. All *S*s, control and experimental, walked the same route and were asked the same questions about the appearance of the scene. At 20-min. intervals *S* returned to the laboratory, where he provided four judgments of equidistance, wearing the MSL. Each experimental *S* had a matched control *S* who was tested with MSL in front of the same eye as his experimental partner. In order to equate the treatments of the control and experimental *S*s, the goggles were removed and replaced before and after each test for all *S*s. The third set of MSL judgments, after 60 min. of exposure, was followed immediately by four judgments with plain glass.

RESULTS

The results reported below are based on the data of 12 experimental and 12 control subjects. Another 18 *S*s were rejected in the initial preadaptation stage of the experiment, for reasons of poor stereoacuity. The rejected subjects complained of the difficulty of the task and exhibited highly variable performance. Their 'plain-glass judgments' often varied over a range that included the values of the ideal MSL effect, and their 'MSL judgments' often were indistinguishable from their plain-glass judgments.

Preexposure plain-glass judgments. The preexposure equidistance settings were highly accurate. The mean error was .56 cm. (SD = .37) and .55 cm. (SD = .46) for the experimental and control groups respectively. In the experimental group, six subjects set the left line a little forward, and six subjects set the right line forward. In the control group, eight subjects set the left line forward and four subjects the right line.

Preexposure MSL judgments. The MSL produces an apparent rotation of the frontal plane. Therefore, when the two test lines are objectively equidistant they appear to be separated in depth, the line on the same side as the MSL appearing farther back. For the viewing distance of 6 ft. and an interpupillary distance of 6.6 cm., Ogle's equation predicts that a 5% MSL will produce an apparent separation of 9.1 cm.[8] Therefore, when a subject is instructed to make an equidistance setting, he should move the line in front of the lens forward by 9.1 cm.

All 24 subjects made adjustments in the predicted direction. Table I shows the mean separation in depth when the two lines were judged to be equidistant. The uncorrected means are the means of the actual settings. The corrected means were derived by comparing each subject's MSL setting with his plain-glass setting. When the plain-glass setting exhibited a bias in the direction of the MSL effect, the MSL setting was corrected by subtracting the initial bias. When the plain-glass setting was in the opposite direction, the MSL setting was corrected by adding the initial bias. The two sets of means differed very little.

Analysis of variance of the corrected data showed only one significant effect. The MSL effect was greater when the lens was in front of the left eye [F (1, 22) = 4.52, $p < .05$]. This difference has been observed by earlier investigators,[9] but it remains unexplained. The experimental and control groups did not differ signifi-

---

[8] The equation for predicting the apparent rotation of the frontal plane is $\tan \theta = [(m\text{-}1)/2m] \cdot [y/a]$, where m is the percent lens magnification, y is the viewing distance, and a is ½ the interpupillary distance.

[9] See Ogle (n. 3).

TABLE I

MEAN PREEXPOSURE MSL EFFECT: SEPARATION IN DEPTH

Group	Left eye		Right eye		Mean of both eyes	
	Uncorr.	Corr.	Uncorr.	Corr.	Uncorr.	Corr.
Experimental	9.78	9.77	7.26	7.27	8.52	8.52
Control	7.98	7.89	7.48	7.58	7.73	7.74
Mean of both groups	8.88	8.83	7.37	7.42	8.12	8.13

Note: All measures in cm.

cantly, nor was the interaction (of eye and experimental or control groups) significant.

Postexposure MSL judgments. The MSL judgments following exposure provide a direct measure of adaptation. If the significance of disparity has been reevaluated, the postexposure MSL effects should be smaller than the initial, preexposure MSL effects. In the case of full adaptation, the MSL judgments would not differ from the preexposure plain-glass judgments.

Table II shows the mean MSL effects prior to the one-hour exposure period and at the three 20-min. postexposure intervals. The combined data for the two eyes show a drop of about 40% after 20 min. of MSL exposure under the experimental condition. There was little change at 40 or 60 min. That this drop was due to MSL exposure is shown by the data for the control subjects, who were treated the same except for the absence of the MSL. The control subjects' combined postexposure MSL effects never differed by more than 6% from their preexposure effects.

The difference between each subject's pre- and postexposure MSL judgments was determined, and an ANOVA computed on the differences. Only the main effect of experimental treatment (experimental versus control) was significant [F (1/20) = 6.58, $p <$ .025]. The mean differences were 3.29, 3.04, and 2.92 at the 20-, 40-, and 60-min. tests for the experimental group, and .07, .40, and .10 at the 20-, 40-, and 60-min. tests for the control group. Thus the experimental group exhibited significantly greater shifts in the adaptive direction than the control group. Six t tests were computed to determine whether the mean shifts differed signifi-

TABLE II

MEAN PRE- AND POSTEXPOSURE MSL EFFECTS

Condition	MSL eye	MSL preexposure	Postexposure (min.) 20	40	60
Experimental	Left	9.77	7.20	6.60	6.89
	Right	7.27	3.25	4.36	4.32
	Combined mean	8.52	5.23	5.48	5.60
Control	Left	7.89	7.79	7.92	8.51
	Right	7.58	7.75	6.76	6.77
	Combined mean	7.74	7.77	7.34	7.64

Note: All measures in cm.

cantly from zero. All three experimental group means were significantly greater than zero ($.02 < p > .002$). None of the control means differed from zero. Thus, the experimental subjects exhibited a significant adaptive shift; the control subjects did not.

Postexposure plain-glass judgments. The postexposure plain-glass judgments measure the aftereffect of adaptation. The rule of thumb for predicting the aftereffect is that when the MSL is removed, a subject will judge two lines to be equidistant when the lines are set to produce the same disparity that was produced by the MSL. The aftereffect will thus be a deviation of judgment of equidistance from true equidistance, with the line on the same side occupied formerly by the MSL set farther back.

Each subject's aftereffect was expressed as a difference between his pre- and postexposure plain-glass judgment. Shifts in the direction of the predicted aftereffect were assigned a positive sign and shifts in the opposite direction were assigned a negative sign. The algebraic mean shift was 3.17 cm. ($SD = 1.68$) for the experimental group and $-.005$ cm. ($SD = .80$) for the control group. The difference was highly significant [$t\ (22) = 7.12$, $p < .022$]. All 12 subjects in the experimental group showed positive aftereffects; 7 control subjects gave small positive shifts and 5 gave small negative shifts.

Another index of the magnitude of the aftereffect is provided by expressing the obtained shift as a proportion of complete adaptation (i.e., of full compensation for the MSL effect). The MSL effect can be specified in two ways: (*a*) the subject's preexposure MSL effect, or (*b*) the ideal geometric effect predicted by Ogle's equation. Table III shows the percent aftereffect computed in these two ways. The control group showed no aftereffect, while the magnitude of the aftereffect for the experimental condition was about 37% of full adaptation.

Correlation between adaptation and aftereffect. The final set of eight judgments was used to assess the rank-order correlation between the direct measure of adptation (MSL on) and the aftereffect measure (plain glass). The Spearman r_s correlation was not significant ($r_s = .21$). Unfortunately, although these two measures were temporally contiguous, they were always secured in the same order, MSL judgments preceding plain-glass judgments.

TABLE III

PERCENT AFTEREFFECT

	% of preexposure MSL effect			% of geometric MSL effect		
	Eye			Eye		
Condition	Left	Right	Combined mean	Left	Right	Combined mean
Experimental	31.22	47.42	39.32	35.27	34.43	34.85
Control	−1.45	2.88	.72	−1.60	1.50	−.05

DISCUSSION

The experiment was designed to show that durative exposure to conflicting information leads to modification of the disparity-depth relationship. The expected modification was presumed to reflect a reevaluation of binocular disparity as a result of exposure to discrepancy. Evidence for the reevaluation was found in the predicted postexposure shifts that were observed in the adaptation and aftereffect measures. The fact that these shifts were observed under testing conditions that eliminated competing depth cues shows that the shifts did not simply reflect ongoing suppression of disparity in favor of other cues. Instead, the shifts seemed to be due to recalibration of the relationship between disparity and perceived depth. The results cannot be attributed to deliberate correction. In the case of the adaptation measure, the subject could not know that the MSL was present. And in the case of the aftereffect, deliberate correction would not lead to shifts. Nor could the shifts be due to uncontrolled factors associated with the exposure period. The control subjects who received the same treatment except for omission of the MSL did not exhibit adaptive shifts.

It should be noted that our interpretation differs from the one that Burian and Ogle favor as an account of Burian's data (n. 5). These authors attribute Burian's data to temporary domination by monocular depth cues. As support for this interpretation, Ogle offers the observation that the stereoscopic depth effects due to the MSL reappear under testing conditions that are free of monocular cues.[10] However, our findings show that the occurrence of adaptive shifts is *not* limited to conditions that provide competing cues. Parenthetically, although it has been overlooked by Ogle and also by Burian, there is evidence in Burian's own data that an adaptive shift can be observed under testing situations that provide little monocular depth information.[11]

The nature of the reevaluation process cannot be settled by this experiment. The following explanation is offered as a tentative hypothesis. When binocular disparity is paired with a competing set of monocular cues a conflict is introduced. The conflict is resolved initially by suppression of the information provided by disparity.[12] Suppression commences almost at once, as evidenced by the fact that even at the outset of MSL exposure, the subject does not report full MSL effects if the visual field contains the normal complement of monocular cues. The suppression process sets the stage for the reevaluation of disparity. The reevaluation is assumed to be an associative process, proceeding along lines similar to the development of stimulus equivalence. That is, by pairing disparity with dominant monocular cues, disparity becomes equivalent to the monocular cues as a stimulus for depth. Experiments are under way to examine this hypothesis by investigating the retention of the new depth-disparity relationship as a function of various interpolated activities that provide different opportunities for unlearning the new relationship or relearning the old one.

---

[10] K. N. Ogle, Special topics in binocular spatial localization, in H. Davson (ed.), *The Eye,* Vol. 4, 1962.

[11] Burian (n. 5); see his Figure 5 on p. 652 and his discussion on pp. 653–654.

[12] The term 'suppression' is being used in a purely descriptive sense. We do not intend to advance any of the specific theoretical connotations that the term may have in the vision literature (e.g., in analyses of binocular rivalry or strabismic amblyopia).

SOME FACTORS DETERMINING EFFICIENCY OF SELECTIVE ATTENTION

CHARLES W. ERIKSEN and JOHN W. ROHRBAUGH,
University of Illinois at Urbana-Champaign

Abstract. Studies of visual memory have used a technique whereby a brief presentation of a display is followed some milliseconds later by a probe designating the position that had been occupied by an element of the display. While the major interest has been in the decreasing accuracy of reporting the element as the indicator is increasingly delayed, an equally important problem is the ability of these indicators to selectively determine attention in a matter of milliseconds. The present experiments determined the efficiency of selective attention as a function of the number of elements in the display and the physical spacing between adjacent elements. The results are interpreted in terms of an attentional focusing process.

The technique introduced by Averbach and Coriell for studying visual information or iconic storage has proven quite fruitful.[1] Essentially it consists of presenting simultaneously an overload of letters or numbers in a brief tachistoscopic exposure and then subsequently presenting a probe, or indicator, that designates one element or subset of elements about which the subject is to report. Even though a large number of letters may be simultaneously presented, if the probe occurs within 100–200 msec. following the display the subject can report with high accuracy the designated letter. The generally accepted interpretation of this result is that while the display as a whole may greatly exceed the subject's capacity to process, he can—if the probe occurs before the icon of the display has decayed beyond intelligibility—use his available processing capacity to selectively process the designated element or subset of elements.

While the technique has been used primarily to study character-

---

Received for publication May 12, 1970. The investigation was supported by United States Public Health Service Research Grant MH-1206 and by United States Public Health Service Research Career Award K6-MH-22,014.

[1] E. Averbach and A. S. Coriell, Short-term memory in vision, *Bell Syst. tech. J.*, 40, 1961, 309–328.

istics of iconic storage and decay, it also offers a method for studying selective attention. The process by which one of a number of equally potent stimuli is selected and processed and the others ignored would appear to be the essence of a concept of attentional selectivity. That the selection process is not instantaneous has been shown by Averbach and Coriell (n. 1). They employed a backward-masking procedure and obtained estimates of over 200 msec. for the relevant item to be selected from the display. Similar durations of the selectivity process have been obtained by Eriksen and Collins, using a somewhat different procedure.[2]

While it seems well established that selective attention requires a finite period of time, little is known about other variables that may affect efficiency (time or accuracy). Previous results have shown that the efficiency is dependent on the number of elements in the display from which the relevant item is to be selected.[3] However, the number of display elements is typically confounded with the spacing between elements. As the number of elements increases, their proximity to each other also increases, which could act to reduce the legibility of each of the elements irrespective of characteristics of the selective process.[4] In Experiment I the efficiency of selective attention was studied as a function of the number of elements in a display for which the spacing between adjacent elements was controlled.

EXPERIMENT I

Method

Subjects. Six students at the University of Illinois (women) served as paid *Ss*. All had normal or corrected-to-normal vision.

Apparatus and stimuli. A three-field Scientific Prototype Model GA tachistoscope equipped with Sylvania F4T5/CWX bulbs was used. All three fields were equated to 10-mL. luminance. A small black cross was exposed in Field I, the adaptation field, which remained illuminated at all times except when the stimulus and the probe were presented in Fields II and III.

---

[2] C. W. Eriksen and J. F. Collins, The temporal course of selective attention, *J. exp. Psychol.*, 80, 1969, 254–261.

[3] C. W. Eriksen and J. Rohrbaugh, Visual masking in multielement displays, *J. exp. Psychol.*, 83, 1970, 147–154.

[4] J. F. Collins, Correlations in sensitivity on different foveal areas, doctoral dissertation, University of Illinois, 1970; M. C. Flom, F. W. Weymouth, and D. Kahneman, Visual resolution and contour interaction, *J. opt. Soc. Amer.*, 53, 1963, 1026–1032.

Stimulus displays were constructed by mounting the four capital letters A, T, V, and H on white plastic cards. The letters were Paratipe No. 11316 pressure letters, each subtending .2° of visual angle. Since three different levels of letter saturation were studied, each of the stimulus-display cards contained either 4, 8, or 12 letters. The arrangement of these different letter saturations on the display depended on the display condition. Both display conditions, distributed and grouped, consisted of an imaginary circle 2.2° of visual angle in diameter centered on the fixation cross. For the distributed displays, the 4-letter saturation presented a letter in the positions corresponding to 2, 4, 8, and 10 o'clock. For the 8-letter saturation, letters occurred at the positions corresponding to 12, 3, 6, and 9 o'clock, and the other four letters were distributed midway between these four points. For the 12-letter saturation, a letter occurred at each of the 12 clock positions.

The grouped-display condition was employed to control for the spacing between adjacent letters as the saturation varied. At each letter saturation, the letters presented always occupied adjacent clock positions. In order to maintain stimulus materials at a manageable number, not all possible arrangements of 4- and 8-letter displays were employed in the grouped condition. Instead, for the 4-letter saturation, the letters occurred in either the positions 12, 1, 2, and 3 o'clock; 3, 4, 5, and 6 o'clock; 6, 7, 8, and 9 o'clock; or 9, 10, 11, and 12 o'clock. For the 8-letter saturation, the arrangements chosen were such that the gap in the arc of stimuli was equally often between 12 and 3 o'clock, 3 and 6 o'clock, 6 and 9 o'clock, and 9 and 12 o'clock. For both distributed and grouped displays, the 12-letter saturation was identical, with the letters occurring in each of the 12 clock positions and thus equally spaced.

For the distributed displays, there were 12 stimulus cards for each of the three saturation levels. The letters were arranged in different random orders on each card, with the constraint that each of the four letters occupy each position equally often for each set of 12 cards. For the grouped displays, a set of 24 cards was constructed for each of the 4- and 8-letter saturation levels. The letters were arranged in a different random order on each card, with the constraint that each of the different letters occupy each clock position equally often.

When the probe preceded or followed the display, it was presented in Field III of the tachistoscope and consisted of a black bar subtending .5° × .1° of angle and located within the circular display so as to constitute a partial radius of the circle pointing to the letter that was being designated. The proximal end of the bar was separated from the target letter by .4°. This probe was presented either 100 msec. before, simultaneously with, or 50 msec. following the display. For simultaneous display-and-probe presentation, a transparent plastic card containing the probe was inserted in front of the display in Field II. When presented before or after the display, the transparent plastic card was placed in front of a blank white stimulus card in Field III and a transparent plastic card without a probe inserted over the stimulus display. Under all conditions the viewing field as observed by *S* was circular, subtending 3.56° of visual angle.

Procedure. Prior to the experimental sessions, all *S*s participated in five practice sessions. During these sessions an exposure duration was established, for each *S*, that yielded approximately 60% identification accuracy on the distributed display with a saturation of 12 letters and simultaneous presentation of the probe. Once established, this duration was then used for that *S* as the display-exposure duration and probe-exposure duration in all conditions of the experiment. Following practice each *S* participated in 12 experimental sessions, 6 under each of the two display conditions. Within each session 12 judgments were obtained for each of the saturation levels under three interstimulus interval (ISI) values; the probe was presented and terminated 100 msec. prior to display onset, the probe was simultaneous with display onset, or the probe followed offset of the display by 50 msec. The presentation order of ISIs and saturation combinations was balanced across sessions and *S*s, as were the display conditions. Under each experimental condition, each of the 4 letters was indicated in each position an equal number of times throughout the experiment. The order of trials was randomized by shuffling a set of 12 stimulus cards for each block of 12 trials. The average exposure duration of the stimulus displays and the probe across *S*s was 10.3 msec., with a range from 7 to 14 msec.

Before each trial *S*s were instructed to fixate the cross and, when it appeared in good focus, to trigger stimulus onset by operating a hand switch. They were required to specify on each trial which of the four letters had been indicated by the bar probe. Each experimental session was preceded by a period of light-adaptation and 12 warm-up trials with the 12-item display and simultaneous presentation of the probe.

Results

In Figure 1 percent correct identifications of the letter indicated by the probe are shown for the distributed and grouped displays for the three different letter saturations. The results are shown as a function of the ISI between probe and display. For the grouped condition, the data have been omitted for those trials when the probe designated the letter in an end position of the group. Since these end letters have a close adjacent letter on only one side, inclusion of the trials where they were the target letter would tend to obscure the effects of close spacing of adjacent letters.[5]

One of the more prominent effects shown in Figure 1 is that for letter saturation. For both the distributed and the grouped displays, performance was superior for the 4-letter saturation and became progressively worse as the saturation increased to 8 and to 12 letters [$F\ (2, 10) = 18.3,\ p < .01$]. The effect of spacing

---

[5] Averaged through ISIs and saturation, the mean number of errors made when the end-position letters were probed was 27%, as opposed to a mean of 34% for inner-position letters. However, an analysis of variance (conditions, ISIs, and subjects) that included the end-position letters in the grouped condition yielded the same significant effects as were obtained when these data were omitted.

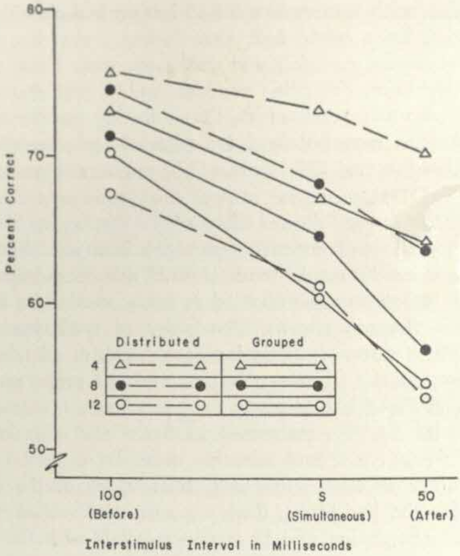


FIG. 1. Percent correct target identifications as a function of ISI, display saturation, and grouped versus spaced conditions

between adjacent letters was also apparent. For both 4-letter and 8-letter saturations, performance was better on the distributed display than on the grouped display [F (1, 5) = 7.7, $p < .05$]. Since the stimulus conditions were identical for the 12-letter saturation of both the distributed and grouped displays, no differences would be anticipated here except for sampling variability.

The effect of ISI between probe and display [F (2, 10) = 27.3, $p < .01$] is consistent with the previous findings of Eriksen and Collins (n. 2). It apparently reflects the processing time required to apprehend the probe and direct attention to the relevant item in the display, during which time decay of the display icon is occurring. That the effect of ISI was uniform for both display conditions and all three saturation levels is indicated by the lack of significant interactions of ISI with these two variables.

There is the possibility that the effect of letter saturation may in part have been artifactual, particularly for the 4-letter saturation on the distributed display. In order to keep the number of displays at a manageable number, the letters under this condition occurred only in positions corresponding approximately to 2, 4, 8, and 10 o'clock. If these positions in the visual field were superior to the other eight clock positions, performance for the 4-letter displays, and to a certain extent for the 8-letter displays, might be artificially high.

To evaluate this possibility, the data were rescored for the 8-letter and 12-letter saturations for both distributed and grouped conditions, employing only the data obtained when the probe designated a letter that corresponded to the 2, 4, 8, or 10 o'clock position. As can be seen in Table I, this reanalysis had little effect on the previous conclusion as to the saturation effect. Here mean percent correct identifications are shown for all three letter saturations when all target positions were scored, as opposed to only targets in positions corresponding to 2, 4, 8, and 10 o'clock. The data in the table are broken down further by the distributed and grouped conditions.

TABLE I

MEAN PERCENT CORRECT TARGET IDENTIFICATIONS AS A FUNCTION OF CONDITIONS AND LOCATION OF TARGETS IN THE DISPLAY

Location and condition	Letter saturation 4	8	12
All positions			
Distributed	73	69	61
Grouped	68	64	62
2, 4, 8, and 10 o'clock positions			
Distributed	73	71	63
Grouped	68	64	62

The sources of errors in identifying the target letter were examined further by determining the extent to which subjects tended to confuse the probed letter with an immediately adjacent letter. Periodic checks throughout the experiment assured that the subjects could always report correctly the clock position that was probed. Nonetheless they could have been uncertain as to which of several letters occupied that position. If they were confusing an adjacent letter with the target, then it would be anticipated that their responses would match letters immediately adjacent to the target more often than chance would require.

The data for the 8-letter saturations in the distributed and grouped conditions were analyzed for the frequency with which the subjects' responses would have been correct for the letter in the adjacent positions on each side of the probed letter. Chance expectancy was determined by the formula

$$C = P(T = A) + (1 - P)(T \neq A)\ 1/3,$$

where C equals chance expectancy, P is percent correct identifications, $(T = A)$ is the number of occasions in the set of stimuli where target and immediately adjacent letters were the same, and $(T \neq A)$ is the number of occasions in the set of stimuli where the target and immediately adjacent letter were not the same.

For both distributed and grouped conditions, the frequency with which the subjects' responses matched the immediately adjacent letter was significantly greater than chance expectancy ($p < .01$) when evaluated via a binomial test. In Figure 2 the percent difference over chance expectancy by which responses matched the immediately adjacent letter is shown for the distributed and grouped conditions as a function of the ISI between probe and display. The effect was greater for the grouped condition but not significantly so ($F < 1$). The ISI variable, however, was significant [$F\ (2, 10) = 4.54$, $p < .05$]. For both conditions there was little tendency to report an immediately adjacent letter if the probe occurred and terminated 100 msec. before the display was presented. However, if the probe occurred with the display, the number of confusions of target letter with adjacent letters increased, and the number of such confusions became even greater when the probe was delayed 50 msec. after display termination.

The possibility that the ISI effect shown in Figure 2 was in part a function of a 'ceiling effect' needs to be considered, since fewer errors were made at 100 msec. before ISI than at 50 msec. after. It seems unlikely that a ceiling effect would affect the data, since even at the before-ISI value, errors were over 25%. Nevertheless, the frequency with which the subjects' responses matched the letter *two* positions removed from the target was determined. Any ceiling effect should affect this measure in the same way that it would affect the measure for the immediately adjacent letter. That a ceiling effect is not sufficient to account for the ISI result in Figure 2 is shown by the finding that matches between the subjects' responses and the letter two positions removed were .8%, .7%, and .6% below chance expectancy for the three ISI values. A t test on the data obtained at 50 msec. after ISI, between the immediately adjacent

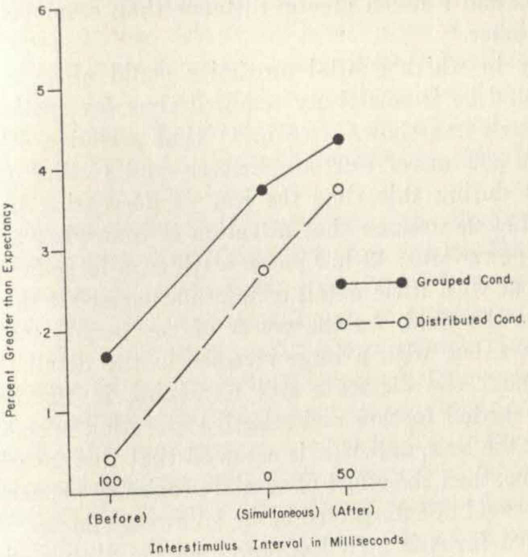


FIG. 2. Percentage by which responses matched the immediately adjacent letter, corrected for chance expectancy, as a function of ISI and conditions

and two-stimuli-removed positions, was significant at the .05 level on a one-tailed test.

Discussion

The ability of a subject to report the probed element in a briefly presented multielement display is a function of both the number of elements in the display and their spatial proximity. The spatial-proximity effect may be independent of a selective attention process. The close spacing of adjacent letters might impair the legibility of the target letter much in the manner that visual masking occurs through contour inhibition or interference. Flom, Weymouth, and Kahneman (n. 4) have shown that the ability to detect the location of the critical detail in a Landolt C is impaired if black bars are located within about ⅓° of arc on each of the four sides of the C. If such an effect were operating in the present experiment, it would reduce the legibility of the letters independent of any selective attentional process. However, the closest spacing between letters was nearly ½° of angle, beyond the effective range found by

Flom et al., and a much greater distance than employed in visual-masking studies.

One way in which spatial proximity could affect an attention process would be if selectivity required time for spatial precision. Prior research has shown (see n. 1) that selective attention can require 200–300 msec. Further, Eriksen and Collins (n. 2) have found that during this time the icon is decreasing in clarity or legibility. Let us assume that attention is somewhat analogous to a variable-power lens. At low power settings, a large field of view is obtained, but with little detail or information about the individual elements in the field. As the power of the lens is increased, the field narrows, but with a large increase in the detail that can be resolved about the elements still remaining in the field. If the analogy is carried further and selective attention is compared to a focusing of the lens, and if it is assumed that this focusing process requires time, then the effect of spatial proximity between letters in the display could be interpreted in the following manner.

Attentional focus is at a low power setting with a wide field of view until the probe is presented. When the probe occurs, the attentional system begins to focus on the probed location, but during the time the focusing is occurring, the icon of the display is decaying. If there is insufficient time—before the display decays beyond intelligibility—for attentional focus to reduce the field to encompass only the probed element, two or more letters may still be in the attentional field. This could lead to confusion as to which of these letters was the probed element. Accuracy in identifying these several alternatives could also be expected to be less, since the power of the attentional field has not had time to increase to a maximum resolution. This lack of sufficient time for complete attentional focusing would thus result in errors from two sources: confusion as to which of several elements was the probed one, and less resolution or identification accuracy for the elements remaining in the attentional field.

The data in Figure 2 are commensurate with such a model. These data indicate that confusions did occur with immediately adjacent letters, but more important, they also show that the tendency for these confusions increased with decreasing time for attentional focusing. If the indicator preceded and terminated 100 msec. before the display, not only was performance in identifying the target letter better, but there was virtually no tendency to confuse an immediately adjacent letter with the target. When there was less time for attentional focusing, as when the probe was simultaneous with the display or was delayed 50 msec. after it, adjacent-letter confusions became more frequent and identification accuracy of the target letter decreased. This result is consistent with our expectations (outlined above) for the event that the attentional field did not have time to focus to the maximum before the display icon decayed.

Letter saturation of the display was found to affect performance independently of spatial proximity between letters. The means by which the number of elements in a display affects the selective process is not clear, but there is evidence that information is processed at a number of different levels prior to specific identification.[6] It may be that all items in the attentional field are processed at least at some minimal level, which may amount to no more than a subject's noting that 'something' was present in different spatial locations. Support for this possibility is found in the observation that our subjects could almost invariably report whether a given display had contained 4, 8, or 12 elements. They had obtained this information despite lack of instruction or set to do so. Also, it was not uncommon for them to report that they had identified one or two letters other than the probed letter in other parts of the field. This information processing, which was extraneous to the task, may have acted to delay the focusing of attention on the probed element or to utilize processing capacity that otherwise would have been available for the directed task.

EXPERIMENT II

How spatial separation between adjacent letters affects attentional selectivity was not entirely elucidated by Experiment I. While there was evidence that the subjects tended to confuse immediately adjacent letters with the target letter, there was not a significantly greater tendency for these confusions to occur in the grouped as opposed to the distributed condition, as might have been expected. Also, in addition to confusions, contour interactions between adjacent letters under the close-spacing conditions may have served to reduce letter legibility. Experiment II was therefore carried out to determine the extent to which contour interactions and interference between adjacent letters played a role in the spacing effect obtained in Experiment I.

---

[6] U. Neisser, *Cognitive Psychology*, 1967.

In Experiment II only three-element displays were employed, with the target always the center of the three. In each display the three elements were separated by one of three distances between their nearest boundaries. In addition to a condition in which all three elements were letters, there was a condition in which the two flanking elements of the target letter were solid discs and another condition in which the flanking elements were Xs. These latter two conditions would be expected to produce contour-interference effects but avoid the possibility of confusions.

Method

Subjects. Four University of Illinois students (two women) with normal or corrected-to-normal vision served as paid volunteers. None had served in Experiment I.

Stimulus materials and apparatus. As in Experiment I *S*s were required to identify a target (A, T, V, or H) on each exposure. The target letter was always the center item in a three-item display and was designated by a concurrently appearing probe. The adjacent items were either Xs, solid black discs, or two of the four letter stimuli. The letter stimuli and probe dimensions were identical to those for Experiment I, with the target letter appearing in a position corresponding to either 12, 3, 6, or 9 o'clock on the imaginary display circle. The adjacent items lay on the same imaginary circle and were separated from the target by either .78°, .38°, or .08° of visual angle. Target letters, discs, and Xs were all the same size on their maximum dimension.

For those display conditions in which the adjacent items were letters, 6 of the possible 16 arrangements of the four letters were selected such that each letter was paired with each of the remaining three.

Procedure. All *S*s served in three practice sessions during which was established an exposure duration that yielded approximately 85% accuracy in identifying the target letter on displays on which only the target letter and accompanying probe were presented. Average exposure duration was 13.5 msec. with a range of 11-17. Practice sessions were followed by six experimental sessions during which each *S* gave 96 judgments for each of the nine spacing-by-type-of-display conditions and an additional 96 judgments for the control condition, for which the target letter appeared only with the probe. Each session contained 10 blocks of 16 trials, each block under a different one of the experimental or control conditions. A block of trials followed one of 12 orders, which were random except for the constraint that each target letter appear equally often in each of the four positions. The stimulus-presentation orders and the conditions were balanced as fully as possible within and across *S*s. Each session was preceded by 16 warm-up trials under the condition for which the target letter was flanked by discs at the .38° spacing.

Results

Figure 3 shows mean percent errors in identifying the target, as a function of spacing between the target and the flanking stimuli, for each of the three types of display (letters, discs, and Xs). Also shown is the performance obtained when only the letter and probe were presented.

Analysis of variance yielded significant effects for the spacing variable [F (2, 6) = 22.9, $p < .01$] and for subjects [F (3, 12) = 7.7, $p < .01$]. The difference between display types did not approach significance ($F < 1$).

Discussion

The data in Figure 3 show that the spacing between the boundaries

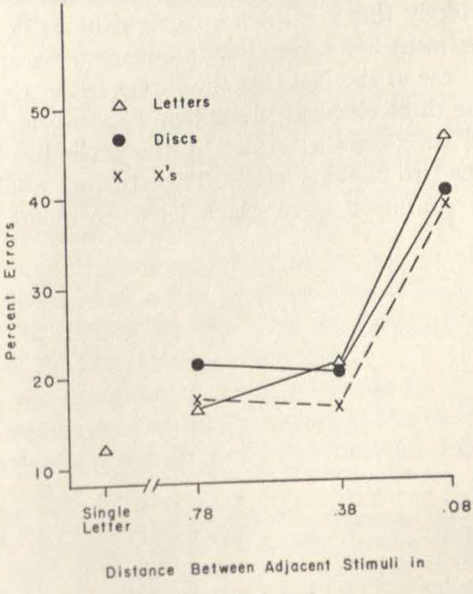


FIG. 3. Mean percent errors in target identification for each display type as a function of spacing between target and flanking stimuli

of adjacent stimuli had an effect on target identifiability only when the distance was .08° of angle. There was no significant or suggestive difference between the .78° and the .38° spacing, irrespective of whether the adjacent stimuli consisted of letters, discs, or Xs. Since the .38° spacing was the distance between boundaries for the closest spacing in Experiment I, the present finding would tend to rule out contour-interaction effects as the basis for the difference obtained between the distributed and grouped conditions of Experiment I. To this extent then, the present results support a conclusion that the spacing effect obtained in Experiment 1 was attributable to a greater tendency for subjects to confuse the target letter with an immediately adjacent one.

The failure to find a significant difference between types of display does not contribute to a confusion interpretation, however. As would be expected from a confusion hypothesis, more errors were made on the letter display at the .08° spacing than were made on the disc or X displays. This would be expected, since with the letter displays, confusions would also contribute to errors over and above contour-interaction effects. However, the difference between the letter display and the other two display types did not approach significance. It is likely that confusion errors on the letter displays in the present experiment had a very low frequency, even under the closest spacing, because of the fact that the target letter was always in the middle of the three elements presented. The subjects may very well have located the target letter not by the probe but by its position relative to the two flanking letters. This strategy would seem to lead to minimum confusion as to which letter to report.

WORD FREQUENCY AND SHORT-TERM RECOGNITION MEMORY

Benton J. Underwood and Joel S. Freund,
Northwestern University

Abstract. In Experiment I, high-frequency words resulted in poorer recognition performance than did low-frequency words only when the 'new' words on the test were also high-frequency words. When they were low-frequency words, recognition was nearly errorless. These findings are interpreted by assuming that a frequency attribute dominates recognition performance and that frequency information for high-frequency lists was influenced by implicit associational responses. Since performance on the recognition test deteriorated somewhat during testing, Experiment II used retention intervals of up to 6 min. and showed that this effect for low-frequency words was not due to the testing itself but to forgetting over time; for high-frequency words, the results were less clearcut. The evidence suggests that forgetting may be retarded by increased temporal discrimination between 'old' and 'new' words.

A number of studies have shown that word frequency, as determined by Thorndike and Lorge, is related to recognition; Gorman's study is an example.[1] More particularly, recognition for words of high frequency is poorer than recognition for words of low frequency. However, a theoretical analysis of recognition leads to expectations that are not in agreement with this finding for all situations. Experiment I, below, was a test of these expectations, which may be understood by indicating the type of recognition test used and the nature of the theoretical assumptions.

The subject was presented a list of 50 words for a single study trial, with full instructions about the nature of the test. On the test, each of the words presented for study ('old' words) was paired with

---

Received for publication March 6, 1970. This work was supported by Contract N00014-67-A-0356-0010, Project NR 154-057, between Northwestern University and the Office of Naval Research. Reproduction in whole or in part is permitted for any purpose by the United States Government.

[1] E. L. Thorndike and I. Lorge, *The Teacher's Word Book of 30,000 Words*, 1944; A. M. Gorman, Recognition memory for nouns as a function of abstractness and frequency, *J. exp. Psychol.*, 61, 1961, 23–29.

a word not presented for study ('new' words) and the subject was asked to designate the old member of each pair. If only the extremes of the frequency dimension are considered, four conditions are required to determine the effects, with such a recognition test, of the frequency of both the old and the new words; for when the old words are high-frequency words, the new words are either high- or low-frequency words, and the same is true when the old words are low-frequency words. Next, the theoretical assumptions, which were:

1. The major attribute of memory involved in discriminating between old and new words is the frequency attribute. Old words have a situational frequency of at least one, new words a situational frequency of zero. It is assumed that the subject's decision for each pair is made by choosing the member of each pair having the highest frequency.
2. Situational frequency does not sum or interact with linguistic frequency (i.e., with Thorndike-Lorge frequency).
3. When a high-frequency word is presented for study, there is a high probability that it will elicit one or more implicit associational responses (IARs), and such IARs also have a situational frequency of one. The IARs to high-frequency words are likely to be other high-frequency words.
4. When a low-frequency word is presented for study, the probability of producing an IAR is low, and if one is produced it is likely to have a higher linguistic frequency than the word producing it. (Points 3 and 4 are essentially statements of known facts: a sample of high-frequency words has a far greater number of interword associations than does a sample of low-frequency words.)[2]
5. The frequency resulting from the actual presentation of a word for study may be augmented by the occurrence of an IAR which is that word.

When the assumptions described were used to predict the outcome of the experiment described, the critical predictions involved the old high-frequency words. During the study trial, IARs are produced by these words. Some of these IARs are other old words, some are new words (other high-frequency words used on the recognition test). When an IAR is another old word, the frequency of this word is augmented and the discrimination on the test should be facilitated. If the IAR to a high-frequency old word is a new word, the so-called old and new words when paired on the test are, essentially, two old words, each with a frequency of one. Therefore, discrimination should be impossible if based on a frequency discrimination. These considerations led to two quite different predictions for the recognition of high-frequency words—predictions which are determined by the frequency of the new words.

---

[2] J. Deese, Frequency of usage and number of words in free recall: The role of association, *Psychol. Rep.*, 7, 1960, 337–344.

When high-frequency words are used as new words, recognition should be depressed. There are two contrary processes operating. First, there is the negative effect of a new word being an IAR during study, which, as noted above, should severely depress performance for such pairs. If an IAR is another old word, and if it is paired with a true new word on the test, performance should be enhanced. However, adding an additional frequency unit to an old word produces a relatively small increase in discriminability.[3] This is to say that a discrimination based upon a 2:0 frequency difference is only slightly better than one based on a 1:0 difference. Therefore, the negative effect should be greater than the positive effect when high-frequency words are presented for study and the new words are also high-frequency words.

But when high-frequency words are presented for study and the new words are low-frequency words, recognition performance should be enhanced, for this situation contains only the positive effect produced by IARs during the study trial. On the test the subject is faced with a series of decisions in which the frequency difference is at least one between the old and new words, and insofar as IARs on the study trial were other old words, the decision may be based on a difference that is greater than one. (When low-frequency words are presented for study, no consideration needs to be given to IARs, since it is assumed that the few IARs produced by low-frequency words are not other low-frequency words and since it is highly unlikely that an IAR would be an high-frequency word if the samples of low- and high-frequency words come from the extremes of the frequency dimension.)

In sum, three levels of performance were predicted: poorest performance under Condition HF–HF (study high-frequency words and test with high-frequency words), best performance under Condition HF–LF, with the other two conditions (LF–HF and LF–LF) falling in between. It can be seen that for Conditions HF–HF and LF–LF the expected outcome was the same as that noted at the outset as having been reported by other investigators, namely, that recognition for high-frequency words is poorer than for low-frequency words. However, according to the theoretical analysis made here, this generalization was not predicted to hold for Conditions HF–LF versus LF–HF.

---

[3] D. L. Hintzman, Apparent frequency as a function of frequency and the spacing of repetitions, *J. exp. Psychol.*, 80, 1969, 139–145.

EXPERIMENT I

Method

Materials. For an earlier study, a random pool of two-syllable AA words was developed, as was an equivalent pool of words with a frequency of one per million.[4] For the present study, 100 HF words and 100 LF words were drawn randomly from the pools; 50 were then chosen randomly from each frequency level to serve as study (old) words, and 50 as new words. The order of the 50 words for study was random but was exactly the same for the two HF lists and the same for the two LF lists. The pairing on the recognition test was random. The order of the pairs on the test was also random, but was exactly the same for the two HF study lists and for the two LF study lists.

Procedure. The study list was shown once to *S* at a 1-sec. rate following complete instructions concerning the nature of the test. After the last word was shown, the instructions for the test procedure were repeated; approximately 30 sec. elapsed from the presentation of the last word to the presentation of the first test pair. The pairs were presented at a 2-sec. rate and *S* was instructed to make a choice for each pair by speaking the word he thought was correct (the old word) during the 2-sec. interval. The instruction that *S* must respond to each pair within the time allotted was strongly emphasized. In the few cases in which an *S* failed to respond to a pair, *E* re-presented the pair (or pairs) after the last pair had been shown.

Subjects. Thirty *S*s were assigned to each of the four conditions, the assignment being made by a schedule of block randomization. The *S*s were college students enrolled in the elementary psychology course.

Results and Discussion

The number of errors was used as the response measure, and the percentages for each of the conditions are shown in Figure 1.

---

[4] The study cited is by B. J. Underwood and B. R. Ekstrand, Word frequency and accumulative proactive inhibition, *J. exp. Psychol.*, 74, 1967, 193–198; see also Thorndike and Lorge (n. 1 above).

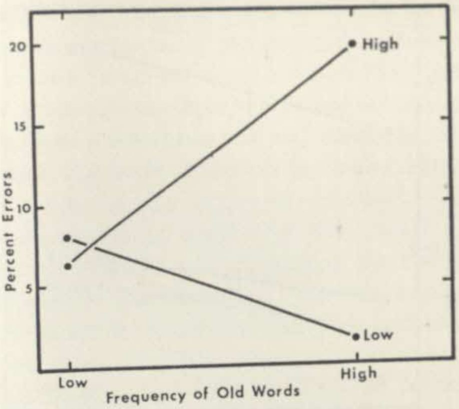


FIG. 1. Recognition errors as a function of the frequency of old words with the frequency of new words as the parameter

It may be noted first that, in accord with previous findings, Condition HF–HF resulted in poorer performance than did Condition LF–LF. However, it can also be seen that performance was nearly errorless under Condition HF–LF. Of the 30 subjects, 20 made no errors, 5 made one error, 2 made two errors, with the remaining 3 subjects committing three, four, and eight errors. Statistically speaking, frequency of old words ($F = 8.69$), frequency of new words ($F = 51.73$), and the interaction ($F = 78.21$) were all significant beyond the .01 level. Using the within-groups variance to derive an error term, the difference between Conditions HF–LF and LF–HF was also found to be significant ($t = 3.00$).

The results are as expected from the theoretical analysis. An alternative interpretation might suggest that in Condition HF–LF performance was determined by the subject's simply choosing the high-frequency word in each pair. There are two facts which oppose this interpretation. If Thorndike-Lorge frequency determined choices in Condition HF–LF, it should also determine equally the choices in Condition LF–HF (the subject would choose the low-frequency word in each pair). But, as noted earlier, these two conditions differed significantly. Still, it might be argued that it took the subject longer to discover the 'rule' in Condition LF–HF than in Condition HF–LF. If this were true, performance in Condition LF–HF should improve during the process of testing. Errors on the first 25 pairs were contrasted with the errors on pairs 26–50

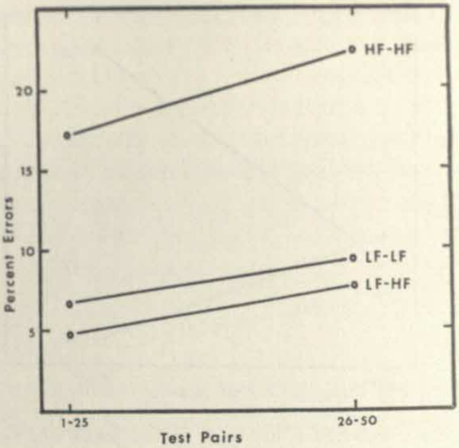


FIG. 2. Changes in recognition errors during testing

(see Figure 2). For reasons which will become apparent shortly, two other conditions are also plotted, but errors were too infrequent to include Condition HF–LF. It is quite clear that recognition did not improve over pairs for Condition LF–HF. In fact, there was a decrement during testing, and the amount of decrement was essentially equivalent for all three conditions, the F for the interaction term being less than one. The decrement during testing was reliable [F (1, 87) = 7.26, $p < .01$].

It has been assumed that IARs are produced during the study trial but no mention has been made of the possibility that they might also occur during the testing phase. Assuming that high-frequency words would be more likely to produce IARs than would low-frequency words, the equivalence of the slopes in Figure 2 indicates that IAR effects were minimal during testing. The question may be raised, however, as to the cause of the decrement during testing. Is it the process of testing per se, or is it forgetting that would occur over time even without the testing? Experiment II provides evidence relevant to the answer.

EXPERIMENT II

It has been assumed that the dominant attribute mediating recognition is the frequency attribute. However, it can be seen that a temporal attribute might support the frequency difference. The use of the terms 'old' and 'new' to identify the two classes of items implies a temporal attribute. If the temporal attribute is involved in the discrimination, a rather unusual prediction emerges. It must be presumed that the discriminative power of the temporal attribute would increase as the retention interval increases. The longer the retention interval, the newer the new items should appear relative to the old. But, of course, the longer the retention interval becomes, the greater the amount of forgetting that could be anticipated from other sources (e.g., loss of frequency discrimination). Thus, while an increase in discrimination via the temporal attribute might be expected as the retention interval increases, a loss of discrimination from other attributes may occur. In spite of the inability to make a prediction, it seemed worthwhile to explore the forgetting curve for recognition over relatively short intervals to see if any signs of a positive effect due to increased temporal discrimination could be detected.

Method

Conditions HF-HF and LF-LF of Experiment I were used, each list being tested for retention after 30 sec., 3 min., or 6 min. This required six groups. The *Ss* were assigned to the block-randomized schedule until 25 were contained in each of the groups. All procedures were exactly the same as in Experiment I. The *Ss* assigned to the 3- and 6-min. conditions were given a symbol-cancellation test during the retention interval to minimize rehearsal.

Results and Discussion

The error percentages are plotted in Figure 3. The retention interval is given in seconds and the points are plotted to include one-half the interval required for testing. During the test there were 50 pairs presented at a 2-sec. rate. Therefore, the 30-sec. retention interval is plotted as 30 + 50 sec., or 80 sec. Similarly, the other two intervals become 230 and 410 sec. The short dotted lines represent the change in performance during the process of testing. The left point for each of these lines is based on items 1–25, the right point on items 26–50, and these are plotted appropriately for the time interval. That is, the left point is plotted at the midpoint of the time required for testing items 1–25 (i.e., 25 sec.), the right point as the midpoint of the time for items 26–50.

It may first be noted that, as in Experiment I, word frequency

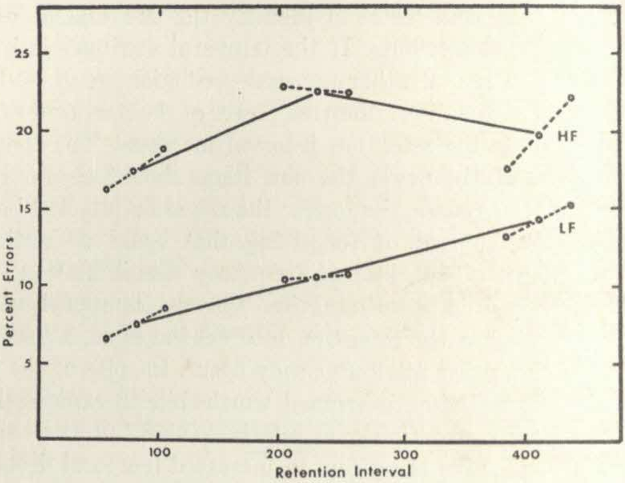


FIG. 3. Recognition as a function of length of retention interval for high-frequency (HF) and low-frequency (LF) lists (solid lines). The dotted lines represent changes in recognition errors during testing.

was a highly relevant variable (F = 39.02). Statistically there is also evidence for forgetting [F (2, 144) = 3.94, $p < .05$], but it is apparent that the loss was not great. The interaction between interval and frequency was not significant (F = 1.62). Overall there was a small testing effect [F (1, 144) = 4.79, $p < .05$], but it did not interact with any of the other variables.

For Condition LF–LF an average testing effect for the three subgroups produced a line with a slope that was almost identical to the slope of the curve over the three retention intervals. This probably means, therefore, that for the low-frequency lists the testing effect was not due to the process of testing as such, but simply represented a loss that would have occurred over time without the testing. The results on this same matter for Condition HF–HF were somewhat ambiguous. An average slope for the three dotted lines would be greater than any reasonable slope assigned the overall retention curve. This would indicate a testing effect per se. Yet, for the 3-min. subgroup there was no testing effect, while for the 6-min. subgroup a very sharp effect was apparent. Furthermore, if only the high-frequency lists are tested, the effect of retention interval was not significant (F = 1.46), indicating that statistically the slope could be zero. No firm conclusions can be reached about the cause of the testing effect for Condition HF–HF.

Is there evidence for increased temporal discrimination as the retention interval increases? Clearly, no such evidence exists for the low-frequency lists. Forgetting is a linear function of time. This does not mean, of course, that a temporal attribute was not influencing the performance; it simply means that for the low-frequency lists there was no evidence that changes in performance over time were being influenced differentially by positive and negative factors. On the other hand, the results for the high-frequency lists do suggest the presence of a process which, at least, retards forgetting. It is as if the discrimination due to the temporal attribute increases at a more rapid rate over time than does the loss of discrimination produced by other processes. At best, however, the results on the matter of increased temporal discrimination over the retention intervals can only be viewed as suggestive–suggestive of the need to work out techniques that would provide an independent measure of the (presumed) increased discrimination.

EYE-DEPENDENT AND DISPARITY ADAPTATION TO OPPOSITE VISUAL-FIELD ROTATIONS

ARIEN MACK, *New School for Social Research*, and
DEANNA CHITAYAT, *Queens College, City University of New York*

Abstract. The experiment reported demonstrates eye-dependent perceptual adaptation and adaptation to altered retinal disparity in the absence of cyclotorsion. Experimental observers were exposed to a binocular prism system that produced 5°, opposite rotations to the two eyes. Tests of monocularly perceived vertical and of binocularly perceived stereovertical were administered before and after exposure. Results indicate simultaneous and opposite eye-dependent changes in the monocular perception of vertical, and adaptive binocular changes in the perception of stereovertical such that zero disparity evoked the experience of depth. Photographic measurements of cyclotorsion were made during testing to determine whether the adaptive changes were a function of oculomotor changes. No cyclotorsion was found.

In recent years there has been an accumulation of evidence indicating changes in visual perception as a result of exposure to distorting lenses. Evidence of perceptual adaptation to displacement, to rotation, and to distortion of size or shape may be considered evidence of perceptual learning.[1] But while there is evidence that monocular exposure to a consistently imposed visual distortion may result in visual adaptation and that (at least under some conditions) this learning is demonstrated by the eye that has been occluded,[2] there has been very little research on whether the two eyes can adapt independently and therefore can learn to see dif-

---

Received for publication February 9, 1970. The authors gratefully acknowledge the help of Amwar Chitayat, who designed the prism goggles, and Robert Sibley, who designed the testing apparatus. The research reported constitutes the doctoral research of the second author.

[1] W. Epstein, *Varieties of Perceptual Learning*, 1967; E. Gibson, *Principles of Perceptual Learning and Development*, 1969.

[2] S. Ebenholtz, Transfer of adaptation as a function of interpolation of optical tilt to the ipsilateral and contralateral eye, *J. exp. Psychol.*, 73, 1967, 263–267; H. L. Pick, J. C. Hay, and R. H. Willoughby, Interocular transfer of adaptation to prismatic distortion, *Percept. mot. Skills*, 23, 1966, 131–135; J. Foley and K. Miyanshi, Interocular effects in prism adaptation, *Science*, 165, 1969, 311–312.

ferently. Were such eye-dependent perceptual learning possible, the appearance of a visual stimulus that casts identical images on the two retinas could differ, depending on which eye was used to view it.

Foley and Miyanshi (n. 2) have reported limited interocular independence of visual-motor spatial coordination in observers simultaneously exposed to a 20-diopter base-up prism over one eye and a 20-diopter base-down prism over the other eye. (The effect of these prisms is to displace the visual fields of the two eyes in opposite directions, one upward and the other downward.) And Hajos demonstrated eye-dependent adaptation in observers who wore base-left and base-right prisms over left and right eyes respectively for nine days.[3] But because adaptation to displacement has been shown to involve oculomoter changes (changes in actual eye position),[4] it seems reasonable to wonder—particularly in the case of the Hajos findings, which derived from exposure to visual fields laterally displaced in opposite directions—whether the measured adaptive changes are not limited to the capacity of the eyes to diverge. If this were so, it would not be that an object casting identical images on the two eyes has, as a result of learning, come to appear differently. Rather, the changes in eye position alone—the recorrelation of the relationship between eye position and perceived direction—would fully account for the eye-dependent effects.

As a way of investigating the possibility of eye-dependent adaptations, we exposed observers simultaneously to opposite rotations of the right- and left-eye visual fields.[5] We did this on the grounds that there has been no evidence that adaptation to prismatic rotation involved oculomotor changes, and there seemed no good reason to expect that such changes would occur. That oculomotor changes accompany adaptation to displacement seems to follow from the fact that the perception of radial direction is at least in part a function of eye position. There is no evidence that perceived orientation is similarly a function of eye position, nor is there a plausible basis to believe that it is. Therefore it seemed reasonable to assume that adaptation to tilt does not depend on rotation of the eye.[6]

---

[3] A. Hajos, Sensumotorische Koordinations prozesse bei Richtungslokalisation, *Z. exp. angewan. Psychol.*, 15, 1968, 435–361.

[4] S. C. McLaughlin and R. G. Webster, Change in straight-ahead eye position during adaptation to wedge prisms, *Percept. & Psychophys.*, 2, 1967, 36–44.

[5] In order to avoid confusion, the following terms are used consistently: 'vertical' refers to orientations in the frontal plane; 'tilt' refers to rotations in the frontal plane; 'upright,' or 'stereovertical,' refers to orientations in the sagittal plane and describes a line whose top and bottom are equally distant from an observer; 'inclination', refers to rotation in the sagittal plane, a line being inclined when one end is farther away from an observer than the other end.

Interocular transfer. Monocular adaptation to tilt transfers to the eye that was occluded during prism exposure.[7] For eye-dependent tilt adaptation to occur, this interocular transfer would either have to be inhibited or in some way overcome. To determine whether interocular transfer effects had in fact been reduced or overcome in our experiment (or whether the two eyes exposed to opposite visual-field rotations behaved like eyes that had been monocularly exposed to the same amount of rotation for the same period), we tested control observers who were monocularly exposed to the same magnitude and direction of visual-field rotation. These observers, like those exposed to opposite rotations, made *judgments of verticality,* before and after prism exposure, with each eye. Unlike Ebenholtz (n. 2), who tested the occluded eye of an observer exposed to monocular prism tilt only after the exposed eye had been tested, we counterbalanced the order in which the eyes were tested. (Ebenholtz reported an interocular transfer effect of 50%.)

Binocular disparity. Opposite rotations of the visual fields of the two eyes introduce a binocular disparity, which may result in distorted depth perception such that an objectively upright line appears inclined either toward or away from an observer, depending on the direction of the rotations. Therefore, a test for changes in apparent inclination was included in our experimental design. Before and after prism exposure each observer made *judgments of the apparent upright* by adjusting a single luminous line that could rotate either toward or away from the observer until it appeared upright. A difference between before and after judgments in the predicted direction was taken as evidence of adaptation to the distorted disparity. We may note here that there was some reason to expect that such adaptation to distorted disparity might occur. Fender and Julesz reported no change in perceived depth as the disparity between the images to the two eyes was either gradually increased or decreased.[8] Köhler and Emory reported figural aftereffects in the third dimension following the inspection of inclined figures.[9] Wallach, Moore, and Davidson reported a modification of stereoscopic depth after observation through a telesteroscope of rotating wire figures.[10] (This modification was, however, primarily limited to the forms used during training.) Wallach and Karsh reported a decrease in stereoscopic depth perception after 24 hours of monocular viewing.[11] Ogle reported some accommodation to aniseikonic lenses but concluded that his studies were evidence in support of the Hering hypothesis concerning the innate and stable character of corresponding retinal elements.[12]

---

[6] For this reason, we chose to use binocular exposure to opposite rotations of the visual field as a means of investigating eye-dependent perceptual learning.

[7] Ebenholtz (n. 2 above).

Any evidence of the predicted changes in perceived inclination in the present experiment would be evidence that the depth significance of disparity had been modified. If an objectively upright line were seen as upright before binocular exposure to opposite field rotations, and then seen as inclined after exposure, we would have clear evidence of a recalibration of the relationship between disparity and depth. Images with zero disparity would be the basis for perceiving inclination.

In order to determine whether any obtained changes in perceived depth were merely instances of figural aftereffects in the third dimension[13] or if they depended on the sensory information apparently relevant to perceptual adaptation (e.g., information derived from movement),[14] a group of stationary observers was tested. These were not permitted to move while they were exposed to opposite rotations of the visual field; their results were taken as an index of the magnitude of three-dimensional adaptation attributable to configurational adaptation. The difference between their results and those of the subjects who were permitted to move while wearing the binocular prism goggles may also be viewed as the degree to which movement-produced information facilitates adaptation. (The stationary observers also provide information on the two-dimensional figural aftereffects involved in normalization of tilt; see the end of the first section on results.)

---

[8] D. Fender and B. Julesz, Extension of Panum's fusional area in binocular stabilized vision, *J. opt. Soc. Amer.*, 57, 1967, 819–829.

[9] W. Köhler and D. A. Emory, Figural after-effects in the third dimension of visual space, this JOURNAL, 60, 1947, 159–201.

[10] H. Wallach, M. E. Moore, and L. Davidson, Modification of stereoscopic depth-perception, this JOURNAL, 76, 1963, 191–204.

[11] H. Wallach and E. Karsh, The modification of stereoscopic depth-perception and the kinetic depth effect, this JOURNAL 76, 1963, 429–437.

[12] K. Ogle, *Binocular Vision,* 1964.

[13] Köhler and Emory (n. 9 above).

[14] I. Rock, *The Nature of Perceptual Adaptation,* 1966; A. Mack, The role of movement in perceptual adaptation to a tilted retinal image, *Percept. & Psychophys.*, 2, 1967, 65–68.

Oculomotor change. A further problem was whether any obtained changes in binocularly perceived inclination or in monocularly perceived orientation were the result of changes in the position of the eyes. It has been argued that perceptual adaptation may be a function of oculomotor changes.[15] For this argument to apply to adaptation to tilt, it must be argued that an imposed rotation of the visual field results in cyclotorsional movements of the eye, movements which persist and lead to new and atypical resting positions. For example, a clockwise rotation of the visual field may cause the eye to rotate clockwise in an attempt to nullify the imposed tilt. If this rotation were either not registered or came to represent the normal resting position of the eye, then an objectively vertical line, which thus cast its image counterclockwise of the vertical retinal meridian, would appear tilted counterclockwise. Only an objectively clockwise-tilted line whose image fell on the true retinal vertical would appear vertical.

While not very convincing in the case of monocular adaptation to tilt, this argument is more convincing in the case of binocular exposure to opposite tilts. Under this condition, cyclotorsion might in fact occur to minimize disparity and facilitate fusion. Indeed, Ogle (n. 12) speaks of one of the sources of cyclotorsional movements as the "urge to fusion." When there is a disparity between the images, the eyes may rotate in opposite directions to minimize disparity and facilitate fusion. A line inclined either toward or away from an observer thus casts disparate images on the retinas of the two eyes—images which are rotated in opposite directions about the foveas. These counterrotations, if large enough to stimulate noncorresponding retinal directions will, according to Ogle, provide a stimulus both for cyclotorsional eye movements (wheellike counterrotations of the eyes) and for stereoscopic depth perception. Such stimulation will "tend to compel cyclotorsions of the eyes so that the images will fall as near as possible on corresponding retinal points." Ogle continues: "Under circumstances where no actual cyclotorsion can occur the inclination of the line is perceived by stereoscopic perception correctly inclined. If, however, owing to the compulsion for fusion, cyclotorsional movements of the eyes do take place, the images of the line will fall nearer corresponding retinal meridians. The functional rotary deviation between the images, which depends upon the relationship of images to retinae, will then be reduced or even eliminated, and the line would appear normal (i.e., not inclined). As far as the eyes are concerned, the images then fall on retinal meridians corresponding to those of a line perpendicular to the visual plane with no cyclotorsion of the eyes" (p. 104).

---

[15] J. C. Hay and H. J. Pick, Jr., Gaze contingent prism adaptation: Optical and motor factors, *J. exp. Psychol.,* 72, 1966, 640–648.

In order to argue that the adaptation effects obtained in the present study were the result of cyclotorsional eye movements that persisted over time, the combined cyclotorsion of the two eyes would have to be equivalent to the amount of angular deviation between the images in the two eyes produced by that inclined line which an observer judged to be upright after exposure to oppositely rotated visual fields. Since binocular exposure to opposite tilts of the visual field appears to constitute the most appropriate condition for eliciting cyclotorsion and, therefore, for investigating oculomotor change as an explanation of perceptual adaptation to tilt, *measurements of cyclotorsion* were included as part of the experimental design.

METHOD

Conditions. Exposure to prismatic rotation occurred under one of the following three conditions for all *O*s. In Group I (experimental), 24 *O*s wore binocular prism goggles which produced 5°, opposite visual-field rotations to the two eyes. Of these, 12 *S*s were exposed to a 5° clockwise and a 5° counterclockwise rotation to the left and right eyes respectively (I-CWL). The remaining 12 *O*s were exposed to the reverse rotations (I-CWR). In Group II (monocular control), 16 *O*s were monocularly exposed to a 5° visual-field rotation. Of these, 8 were exposed to a 5° clockwise rotation (II-CW), and the remaining 8 *O*s were exposed to a 5° counterclockwise rotation (I-CCW). For each direction of rotation, half of the *O*s were exposed with the right eye while the left eye was occluded; for the remaining *O*s, occluded and exposed eyes were reversed. In Group III (stationary control), 8 *O*s were binocularly exposed to opposite prism rotations. As in Group I, half the *O*s were exposed to a 5° clockwise rotation over the left eye and a 5° counterclockwise rotation over the right eye (III-CWL), and the remaining *O*s were exposed to the reverse conditions (III-CWR). All exposures occurred while *O*s were seated and prevented from moving their heads by means of a rigid bite bar.

Exposure and testing apparatus. The optical system that created visual-field rotation consisted of two pairs of Dove prisms mounted in tandem and set on goggle frames, a pair of prisms in front of each eye. The prism closer to the eye was rotatable and served to disorient the visual field. The prism farther from the eye corrected the right-left reversal produced by the rotating prism. The rotating prism was movable by means of a lever calibrated to an accuracy of .5°. When the lever was set at zero, there was no visual-field rotation. The binocular field through the prisms was 25° × 25°. The prism goggles weighed a total of 2.8 lb. In order to counterbalance this weight and thus help support the goggles, they were attached to the front of a helmet. A rod attached to the back of the helmet held an adjustable counterbalancing weight.

The stimulus used for all testing consisted of an electroluminescent line, 12 × ⅝ in., which rotated either clockwise or counterclockwise in the frontal plane or forward or backward in the sagittal plane. The line was calibrated in both planes to within .5°. Either forward and backward or right and left movements of the line were accomplished by a set of strings given to the *O*. For any test the line could move in only one plane. The *O* sat 35 in. in front of the test line with his head held rigidly in place by a wax-bite chin- and headrest. Two movable narrow ¾-in. vertical strips of wood used for occluding one of the eyes during testing were placed 12.5 in. from the eyes. Thus while the luminous line was occluded for one eye, it was possible for the camera, used to measure cyclotorsion, to photograph both eyes simultaneously.

The camera, a 35-mm. Nikon F with a telephoto angled lens, was placed 20 in. from the *O*'s eyes. The camera was placed at an angle of 20° to the *O*'s head; and both of the *O*'s eyes were in sharp focus. A shutter speed of ⅓ sec., a focal length of 2 in., and an aperture of *f* 22 were used. A flashgun was mounted in front of the *O* several inches below eye level. A blue filter placed in front of the flash enhanced the contrast between blood vessels and sclera.

Procedure. In a pretest of apparent inclination, each *O* sat in a dark room with head held firmly by means of a wax-bite chin- and headrest and he was asked to set the inclined test line (alternately inclined forward and backward), so that it appeared upright—with the upper and lower ends of the line thereby appearing equally distant. Four judgments were made. Approximately 3 sec. after each of the first three adjustments, both eyes were photographed.

In a pretest of two-dimensional orientation, an occluder was set in front of one of the *O*'s eyes and he was instructed to set the tilted test line (alternately tilted clockwise and counterclockwise of vertical), so that it appeared vertical. Photographs of both eyes were taken after each of the first three adjustments. Four adjustments were made. This procedure was repeated using the eye that had been occluded during the initial test. The tests of perceived orientation were administered in a counterbalanced order except that the monocular tests were always consecutive. All testing was done with the naked eye.

Total exposure to prisms was 40 min. for each *O*. The *O*s in Groups I and

II wore the prism goggles while walking through the halls and courtyard of the New School. The *O*s in Group III remained stationary and looked down a hallway. For the *O*s in Groups I and III, the exposure period was interrupted by four testing sessions; there were three testing sessions for *O*s in Group II. Each of the testing sessions involved four judgments and consisted of one of the following tests: a monocular right-eye test of perceived vertical (R); a monocular left-eye test of perceived vertical (L); or two binocular tests of perceived inclination (B). The *O*s in Groups I and III received a B posttest after the first 5 min. of prism exposure. This was followed by another 15 min. of exposure and a second posttest. Between all subsequent tests there was a 10-min. exposure interval. (Exposure intervals between tests allowed us to measure adaptive changes which suffered minimal decay.) The order of the last three posttests was counterbalanced across *O*s with the the exception that the monocular tests were always consecutive. The posttest orders in Groups I (I-CWL and I-CWR) and III (III-CWL and III-CWR) were B R L B, B L R B, B BL R, and B B R L respectively. In order to avoid any unnecessary binocular stimulation during the (monocular) exposure period for Group II, there was only one binocular posttest, which occurred after the full 40 min. of exposure. Since there was no initial test of apparent inclination for *O*s in this group, the first 20 min. of exposure were uninterrupted. Posttest orders in Group II (II-CW and II-CCW) were L R B and R L B respectively.

The method used for measuring cyclotorsion was a modification of the photographic procedure described by Howard and Evans.[16] The principal requirement any photographic measure of torsion must satisfy is that it establish an angular relationship between some clearly detectable referent in the eye and some other fixed point. Unlike Howard and Evans, we established as our fixed referent a line connecting the center points of the irises. (Howard and Evans used a plastic U-shaped form which pressed against the *O*'s face and surrounded his eye.) Since we were measuring possible torsion changes after prism exposure, and thus needed to photograph the eye prior to and following the exposure period, we had to establish a fixed referent that would not be changed by removing and replacing the *O* in the testing position. The justification for using the line connecting the centers of the irises as our fixed referent against which torsion could be assessed derives from the fact that disjunctive vertical movements, the only movements that could affect the relationship between the center of the irises, "are not normally produced voluntarily,"[17] and therefore were not expected to occur in the testing situation. Slight changes in head position were not expected to result in misleading measures of cyclotorsion, which was indicated by a change in the angle between the line drawn from episclerotic blood vessels to the line connecting the irises.

Photographs of the eyes were taken while the *O*'s head was in a fixed position relative to the camera. These photographs were later viewed by means of a rear-projection viewer constructed to facilitate measurements.

---

[16] I. P. Howard and J. A. Evans, The measurement of eye torsion, *Vis. Res.*, 3, 1963, 447–455.

[17] H. Davson, *The Physiology of the Eye*, 1950.

A projector was placed approximately 10 ft. from the viewer so that the image filled a 2 × 3 ft. screen. The structure of the eye was thus sufficiently enlarged to allow for precise measurement. The general principle of measurement was to establish a line between the centers of the irises of the two eyes, which essentially determines head position (θA). The relative position of the visible blood vessels in each eye established the relative rotation (θr) of the vessels about the iris. Cyclotorsion was then determined by comparing the difference between head and blood-vessel angle ($\theta A\ \theta r$) from one testing session to another. The difference is an index of cyclotorsion.

Since three photographs were taken during pretesting, we were able to average the readings of these three photos, thereby minimizing any variability in the initial measurement. Since measurement errors are unavoidable,[18] repeated measurement of the same picture and of the difference between pictures was made. Each pretest photograph was measured three times and the mean angle between head and blood vessels was established. This measurement was then used as the standard against which postexposure photographs were assessed. Each postexposure photograph was measured against the initial reference drawing and, again, three measures of each of the three photographs were made. A mean of these measures was established. The resulting score reflected the difference between the position of the blood vessel relative to the head before and after exposure. This difference is the measure of torsion.

In determining the combined left-right cyclotorsion of the eyes (CLRC; the measurement most directly relevant to reductions in disparity because it affects the functional disparity of the retinal images and consequently the appearance of inclination), we found the algebraic difference between the rotations of each eye. The retinal images of all inclined lines are rotated about the fovea in opposite directions. The position of the images of a line relative to the optical center of the two eyes may be determined geometrically by the equation

$$\tan S = \frac{2a}{b} \tan i,$$

where S is the rotary deviation or disparity between images, i is the inclination of the viewed line, $2a$ is the interpupillary distance, and b is the observation distance.[19] This calculation establishes the geometrical angular rotary deviation of the images. Cyclotorsion reduces this deviation. Thus, the perception of inclination (in the dark) depends, according to Ogle, on the functional deviation, which is defined as the geometrical deviation less the CLRC. In order to explain adaptation to inclination in terms of cyclotorsion, the CLRC would have to be large enough to compensate for the geometrical angular deviation induced by the inclined line that the *O* had come to perceive as upright after prism exposure. And in order to account for adaptive monocular changes in the perceived vertical in terms of cyclotorsion, the rotation of the eye would have to compensate for the tilt of the line perceived as vertical after exposure.

---

[18] Two methods were used in estimating error in measuring cyclotorsion. The first method involved determining the factors contributing to possible errors and then estimating their contribution. This analysis yielded an error estimate of .04°. The second method involved estimating accuracy by determining the variability within the pretest photographs. We assumed that any cyclotorsion that occurred during pretesting should be consistent across a pretest session. Thus any variability resulting from this analysis was attributed to errors of measurement. An analysis of 17 sets of photographs (three photographs per set) yielded a standard deviation of .057°. It may be concluded from the error estimate that the errors were on the order of .06° (standard deviation). Based on a standard error of .04°, there is a 95% chance of detecting ± .08° of cyclotorsion.

[19] Ogle (n. 12).

RESULTS

Eye-dependent adaptation to tilt. Table I presents the data relevant to whether eye-dependent adaptation to (frontal) tilt occurred. An analysis of variance revealed that neither test order, nor which eye was exposed to which direction of visual-field rotation, nor the interaction of these factors affected the results. Thus all data from the clockwise-exposed eyes in Group I were combined and considered as a unit, and all data from all counterclockwise-exposed eyes were similarly considered. Of the 24 observers in Group I who were exposed to 5°, opposite rotations of the visual

TABLE I

EYE-DEPENDENT ADAPTATION TO TILT

Group I—CWL				Group I—CWR			
S	Left eye	Right eye	Diff.	*S*	Left eye	Right eye	Diff.
1	+.75[†]	+.75	0	13	+.8	−1.2	2.0
2	−1.60	+.22	−1.82	14	+1.6	−2.4	4.0
3	−1.43	+1.70	−3.13	15	−.2[†]	−1.0	.8
4	−1.00	+1.6	−2.6	16	+.5	+.7[†]	.2[†]
5	−.56	+2.0	2.56	17	+1.5	−.6	2.1
6	−.88	+3.12	4.06	18	+1.5	−.8	2.3
7	−.90	+.8	1.7	19	+3.8	+1.9[†]	1.9
8	−1.62	−.62[†]	1.0	20	+3.1	−2.2	5.3
9	−.32	+1.13	1.45	21	+1.7	+1.4[†]	.3
10	+.5[†]	+.2	.3[†]	22	−2.0[†]	−4.7	2.7
11	−1.0	−2.5[†]	1.5[†]	23	+.3	−1.2	1.9
12	−1.31	+2.25	3.56	24	−.2[†]	−1.5	1.3
±	−.82	+.89	1.7		+1.1	−.96	2.06

Note: All scores in degrees; + indicates counterclockwise, − indicates clockwise.

[†] Direction of change against prediction.

field, 19 adapted significantly to clockwise rotation and of those 19, 14 simultaneously adapted to counterclockwise rotation. The mean adaptation for the clockwise-exposed eye following simultaneous exposure to clockwise and counterclockwise field rotation was .88° ($t = 3.19$, *df* 23 α .005). Thus a line tilted .88° clockwise of the preexposure vertical appeared to be vertical when viewed by the clockwise-exposed eye. This change represents 18% of complete adaptation to a 5° rotated visual field. Mean adaptation for the counterclockwise-exposed eye following simultaneous exposure to opposite field rotations was .97° ($t = 3.13$, *df* 23 α .005). This represents 19% of complete adaptation. The difference in magnitude of these effects was not significant ($t = .48$, *df* 23). The data indicate that the two eyes can adapt in opposite directions as a result of exposure to opposite visual-field rotations. Following simultaneous exposure to opposite rotations, the mean difference in perceived verticality between the clockwise- and counterclockwise-exposed eyes was 1.85° ($t = 5.79$, *df* 23 α .001). This means that as a result of the exposure conditions the clockwise- (counterclockwise-) exposed eye perceived a line as vertical when it was in fact tilted 1.85° clockwise (counterclockwise) of the line viewed as vertical by the counterclockwise- (clockwise-) exposed eye.

These results may be compared with the results for Group II, in which each observer was (monocularly) exposed to only one direction of rotation (Table II). An analysis of variance performed

TABLE II

ADAPTATION TO TILT

Group I (CWL and CWR)		Group II (CW and CCW)	
Left eye	Right eye	Left eye	Right eye
+2.1	*−1.3*	+3.6	*+1.6*
+1.6	*−.9*	+4.1	*+2.1*
−.9	*−2.2*	+3.6	*+2.0*
−1.8	*−3.6*	+2.9	*+.7*
−1.4	−1.45	*+1.7*	+1.4
−1.1	−3.4	*+1.2*	+1.9
−1.56	−2.2	*+2.7*	0
−2.31	−3.4	*+.5*	+2.6
$\bar{X}$ −.67	−2.31	$\bar{X}$ = +2.53	+1.53

Note: All scores in degrees; italics indicate exposed eye; + indicates counterclockwise, − indicates clockwise.

on the data of Group II indicated that neither test order, nor which eye was exposed, nor the interaction of these factors significantly affected the results. Thus, all data obtained from clockwise- and counterclockwise-exposed eyes were pooled. The difference between the pre- and postexposure judgments of the vertical for the 8 observers in Group II exposed to clockwise rotation of the visual field was 1.79° clockwise ($t = 5.84$, df 7 α .001). For the 8 observers exposed to counterclockwise rotation the mean difference was 1.56° counterclockwise ($t = 5.97$, df 7 α .001). This represents 35% and 31% of complete adaptation respectively. In order to assess the degree of interocular transfer in the monocular condition, the effect obtained when testing with the exposed eye was compared with the effect obtained when testing with the eye that had been occluded during prism exposure. (It should be remembered that the order in which the eyes were tested was counterbalanced.) In the case of clockwise exposure, testing with the eye that had been occluded during exposure yielded a pre- to posttest difference of 1.24°. The difference between the clockwise exposed and occluded eye was not significant ($t = .75$, df 7). In the case of exposure to counterclockwise tilt, the eye that was occluded yielded an aftereffect of 2.51°. In this case the greater effect was obtained with the occluded eye. This difference, however, was not significant (t .174, df 7). Interocular transfer of monocular adaptation to tilt appears to have been quite complete.

A direct statistical comparison between the mean effect obtained in Group I as a result of the clockwise (and simultaneous counterclockwise) exposure to a rotated visual field and the mean effect obtained as a result of monocular clockwise exposure in Group II indicates that this difference was significant at the .05 level for a one-tailed test ($t = 1.84$, df 30). The same comparison between Groups I and II for counterclockwise exposure indicates that the mean differences came close to significance at the .05 level for a one-tailed test ($t = 1.05$, df 30; with 30 df a t of 1.69 is necessary for significance at .05). Thus the effect of exposure to either a clockwise or counterclockwise rotation was somewhat greater when that exposure occurred monocularly and the other eye was occluded than when it occurred binocularly concomitant with exposure to opposite rotation. The fact that approximately complete interocular transfer occurred with monocular exposure provides strong evidence of eye-dependent adaptation with exposure to opposite visual-field rotations. We may assume that the evidence of eye-dependent adaptation obtained represents a change over and above an inhibition of transfer. Thus, if in the monocular exposure conditions one eye adapts to clockwise field rotation and comes to see a line tilted 1.7° clockwise as vertical, the other eye also perceives a line tilted approximately 1.7° clockwise as vertical. In the case of exposure to opposite rotation, the only evidence of transfer was the decrease in the magnitude of the effects as compared to a monocularly exposed observer. Under these conditions one eye comes to see a line tilted clockwise, while the other eye comes to see a line tilted counterclockwise as vertical. The fact that the absolute effects were smaller in the binocular conditions is, most probably, evidence of the interaction of the two eyes, which was most clearly seen in the magnitude of the interocular transfer in the monocular condition.

The occurrence of complete interocular transfer in monocularly exposed observers might seem to argue against eye-dependent learning, since when only one eye was exposed both eyes demonstrated equivalent learning. However, the occurrence of apparent eye-dependent learning in Group I argues against such a view. It may be that the eyes will normally behave and learn conjointly, but are capable of independence. Thus when visual input creates the conditions that make independence appropriate, independence is demonstrated. The data obtained supports some such view.

An examination of the effects obtained in Group III, for which the binocular exposure to opposite field rotations was under stationary conditions, indicates both the degree to which the effects obtained in Groups I and II were the result of movement-produced information and the extent to which these effects might be attributable to configurational aftereffects of the kind studied by Gibson and Radner and by Köhler and Wallach (Table III).[20] Again all data from the clockwise- and counterclockwise-exposed eye were pooled into two groups, a clockwise and counterclockwise group. The mean difference between pre- and posttest judgments of the vertical for the clockwise-exposed eye was .12° ($t = .31$, *df* 7) and for the counterclockwise-exposed eye this difference was .05° ($t = 12$, *df* 7). Neither difference was significant. Thus, we conclude that the changes of apparent two-dimensional orientation for Groups

---

[20] J. J. Gibson and M. Radner, Adaptation, after-effect, and contrast in the perception of tilted lines: I, Quantitative studies, *J. Exp., Psychol.*, 20, 1937, 453–467; W. Köhler and H. Wallach, Figural after-effects; An investigation of visual processes, *Proc. Amer. phil. Soc.*, 88, 1944, 269–357.

TABLE III

STATIONARY ADAPTATION TO OPPOSITE VISUAL-FIELD ROTATIONS

Group III—CWL			Group III—CWR		
S	Left eye	Right eye	*S*	Left eye	Right eye
1	−.5	+.9	5	+.6	−.9
2	+.6[†]	−2.4[†]	6	−.2[†]	+1.3[†]
3	+.9[†]	+1.5	7	−.3[†]	−.8
4	+1.4[†]	−.5[†]	8	+.1	−1.8
$\bar{X}$	+.6[†]	−.06[†]		+.05	−.55

Note: All scores in degrees.
[†] Direction of change nonadaptive.

I and II were not the results of configurational adaptation but a function of movement-produced sensory information.[21]

Adaptation to inclination. An analysis of variance revealed that there was no significant difference in the magnitude of adaptation to inclination, measured after 20 min. of exposure, between the two conditions of exposure in Group I or between the several test orders. Therefore, the data from Group I were pooled. Table IV reports the amount of adaptation to inclination first after 5 and then after 20 min. of exposure for Group I.

No inclination data are reported for Group III, since those observers in this condition who were exposed to a simultaneous 5° counterclockwise and clockwise tilt over their left and right eyes respectively made judgments of inclination prior to prism exposure that were approximately 6° below the mean judgments in all other groups. The other observers in Group III, who were exposed to a 5° clockwise and counterclockwise tilt over their right and left eyes, showed a mean adaptive change in apparent inclination of .25° after 5 min. of exposure and a mean change in nonadaptive direction of .05° after at least 20 min. of exposure. These changes are clearly not significant.[22]

In order to determine whether any changes had occurred in the

---

[21] The absence of any two-dimensional figural aftereffect or normalization of tilt is perhaps not surprising, since simultaneous opposite binocular tilts of the visual field hardly seem ideal conditions for eliciting such changes.

[22] The absence of any three-dimensional figural aftereffect or normalization may seem surprising, but it is consonant with the results of earlier studies in which the observers were permitted to view a natural environment rather than a single line or figure; see Mack (n. 14 above).

TABLE IV

ADAPTATION TO INCLINATION

Group I—CWL			Group I—CWR		
S	After 5 min.	After 20 min.	*S*	After 5 min.	After 20 min.
1	9.3	13.50	13	12.2	21.8
2	−2.0*	4.50	14	4.7	15.0
3	4.5	3.75	15	4.4	10.4
4	9.9	.90	16	6.7	16.7
5	10.9	15.90	17	4.8	10.8
6	5.0	7.87	18	9.5	13.5
7	−2.3	19.70	19	12.6	22.0
8	5.3	9.00	20	11.3	28.7
9	6.25	11.75	21	10.2	14.0
10	11.3	22.50	22	8.7	6.0
11	3.75	−6.0	23	3.6	10.4
12	6.62	3.87	24	.3	9.6
$\bar{X}$	5.71	8.93		7.4	14.9

Note: All scores in degrees.
* Direction of change nonadaptive.

relation between retinal disparity and perceived inclination, an observer's judgments of upright made prior to prism exposure were compared with his judgments after 5 and at least 20 min. of exposure. A change was considered adaptive if it was in the direction of prism inclination. A clockwise rotation over the right eye and counterclockwise rotation over the left eye produced an apparent backward inclination of an objectively upright line. The reverse rotations produced an apparent forward inclination.

After 5 min. of exposure, 22 of the 24 observers showed significant adaptation (t = 17.36, *df* 23, α .001). There was a mean adaptive change in the perceived stereovertical of 6.66°. After 20 min. or more (up to 40 min.) of exposure, 23 of the 24 observers adapted. The measured mean adaptive change was 12.42° (t = 8.40, *df* 23 α .001). The difference between the mean change after 5 and 20+ min. of prism exposure was significant (t = 4.30, *df* 20 α .001). Thus there is evidence of an adaptive change in the relationship between disparity and perceived inclination, so that a line inclined in the direction in which the prisms were inclining appeared upright after prism exposure. There also appears to be evidence that this change grows over time. Because interocular transfer was nearly complete in the monocularly exposed observers, there is no reason to expect a change in perceived inclination. There was none.

Cyclotorsion. The data on cyclotorsion were assessed in terms of whether the torsion that occurred could account for changes in apparent monocular tilt or binocular inclination adaptation. If cyclotorsion were the basis of the eye-dependent adaptation to tilt, then the cyclotorsion would have to be in the direction of the imposed prism rotation. That is, if the left eye was exposed to counterclockwise rotation and evinced an adaptive change in the perception of vertical after that exposure, there would have to be evidence of a counterclockwise rotation of that eye in order for cyclotorsion to be considered the basis of the adaptive change. In Group I the amount of cyclotorsion measured in the eye exposed to clockwise rotation was .03° clockwise ($t = .53$, df 8). Measured in the eye exposed to counterclockwise field rotation, it was .02° clockwise ($t = .45$, df 11), which is a change in the wrong direction. Thus, there was no significant cyclotorsion during the testing of adaptation to tilt. Since the correlation between adaptation to tilt and measured changes in eye position was trivial (the correlation between adaptation to clockwise tilt and cyclotorsion in Group I was .04; between adaptation to counterclockwise tilt and cyclotorsion it was .06), we infer that the changes in eye position that did occur were the result of random eye movements and that the eye-dependent adaptive effects obtained were not a function of cyclotorsion.

The data also indicate that there was no significant cyclotorsion during testing for changes in perceived tilt following monocular exposure to tilt. After monocular exposure to rotation, the observers in Group II subjected to clockwise rotation demonstrated a mean change in eye position of .06° clockwise ($t = .48$, df 3). The observers subjected to counterclockwise rotation in that group demonstrated a mean change in eye position of .09° counterclockwise ($t = .86$, df 4). Thus, cyclotorsion does not appear to have been the source of adaptation to tilt.

In order for cyclotorsion to account for the changes in perceived upright following exposure to simultaneous opposite rotations of the visual fields, not only would the eyes have to rotate in opposite directions, but the sign of the difference in the direction would have to be dictated by the nature of the imposed disparity. That is, if the right eye was exposed to clockwise and the left eye to counterclockwise rotation, the right eye would have to rotate clockwise and the left eye counterclockwise in order for cyclotorsion to account for adaptive changes in perceived inclination following prism exposure. The cyclotorsion data obtained during testing sessions in which the observers set the luminous test line to the apparent upright were analyzed in two ways. First, the data obtained from each eye were examined; and second, the difference between the position of the left and right eye was calculated (CLRC). After 5 min. of exposure the eyes subjected to clockwise rotation showed a mean cyclotorsion of .005° clockwise ($t = .06$, *df* 11). After 5 min. of exposure the eyes subjected to counterclockwise rotation showed a mean cyclotorsion of .05° counterclockwise ($t = 82$, *df* 11). After at least 20 min. of exposure the eyes subjected to clockwise rotation showed a mean cyclotorsion of .08° clockwise ($t = 1.27$, *df* 13). After at least 20 min. of exposure the eyes exposed to counterclockwise rotation showed a mean torsion of .09° clockwise ($t = 1.45$, *df*, 12), a change in the wrong direction.

Since the angular relationship between the position of the eyes is directly relevant to the magnitude of angular rotary deviation between the images and, therefore, is directly relevant to perceived inclination, the change in angular relationship was established by combining the cyclotorsion measured in one eye with that of the other eye (CLRC). This directional difference was considered positive when the difference in direction of rotation was in the predicted direction. Thus for Group I-CWL the left eye was expected to show a greater clockwise change than the right eye, and for Group I-CWR the reverse was expected. After at least 20 min. of exposure there was an angular deviation between the two eyes of .004° in a direction opposite to that predicted by a cyclotorsion explanation of adaptation to inclination ($t = .039$, *df* 12). There was a correlation of .17 ($t = .054$, *df* 11) between the magnitude of angular deviation between left and right eye and the measured adaptation to inclination after 20+ min. of prism exposure. We therefore conclude that cyclotorsional movements of the eye were not the source of adaptation to imposed distortions of inclination.

Indirect evidence for this conclusion is the fact that there was no apparent adaptation to either inclination or (frontal) tilt in the case of the Group III observers, who were exposed to simultaneous opposite tilts while stationary. If cyclotorsional movements were the source of the adaptive changes, either to opposite tilt or to inclination distortions or both, adaptation should have occurred as readily under stationary exposure conditions. The fact that there were no measured adaptive changes obtained for Group III suggests that something other than cyclotorsional movements was the basis for the changes occurring in Groups I and II. There is no reason to assume that for cyclotorsion to occur an observer must be mobile.

DISCUSSION

The results demonstrate a change in monocularly perceived tilt and binocularly perceived inclination after exposure to 5°, opposite rotations of the monocular visual fields. The eye-dependent changes in apparent two-dimensional orientation were smaller than changes in apparent two-dimensional orientation obtained in monocularly exposed observers, but clearly demonstrated eye-dependent perceptual learning.

The changes in the perceived upright were such that after at least 20 min. of binocular prism exposure a line inclined 12.42° from the objective upright appeared upright. From Ogle's formula, given above, we may infer that a line placed directly before an observer and inclined 12.42° produced a geometric angular rotation of .91° between images of that line. Since, according to Ogle, the angular deviation between images of an inclined line is—in the absence of cyclotorsional eye movements—equivalent to the retinal disparity occasioned by that line, and since there was no evidence of cyclotorsion, we may conclude that as a result of exposure to opposite visual-field rotations, a disparity of .91°, which is normally associated with the perception of inclination, was associated with the perception of upright.[23] This represents a significant recalibration of the relationship between disparity and perceived depth, and is consistent with other reports indicating the lability and modifiability of this relationship.

The fact that the relationship between perceived depth and retinal disparity was subject to recalibration raises a question about the concept of corresponding retinal points. Normally an objectively upright line placed directly in front of an observer stimulates corresponding points on the retinas of the two eyes. If this line is luminous and viewed in an otherwise dark room, so that the inference that apparent line orientation is strictly a function of

[23] A disjunctive cyclotorsion of .91° was well within the range of sensitivity of the measurement procedure, which was accurate to within .08°.

retinal image orientation is valid, then the generally accepted explanation of why this line appears upright is that the images to the two eyes fall on corresponding and therefore nondisparate retinal points. Corresponding retinal points are sometimes defined as those points on the retina of one eye which would coincide with the points on the other eye if one retina were superimposed on the other.[24] Given this description, retinally corresponding points are equivalent to physically corresponding points. Yet, if retinal correspondence is based on physical correspondence, how can the set of retinally corresponding points ever be changed? Prior to binocular prism exposure to counterrotated monocular visual fields, a luminous line viewed in the dark and inclined away from the observer appears inclined. The perceived inclination is said to follow from the retinal disparity between the images of that line. The image to the left eye falls clockwise, and on the right eye counterclockwise, of true retinal vertical. Thus, noncorresponding retinal points are stimulated. After exposure to clockwise rotation over the left eye and counterclockwise rotation over the right eye, a line inclined toward the observer—so that the image to the left eye is .91° clockwise of the image to the right eye—looks upright. Physically disparate retinal points now correspond. This being the case, the equating of corresponding retinal points with physically corresponding points must be revised. There must be some other criterion that defines correspondence.

Ogle's definition (n. 11) seems more appropriate. He defines corresponding retinal points as points which give rise to the "same primary visual direction," or "primary subjective direction" (p. 103). This definition seems to be consistent with the data indicating a recalibration of the depth-disparity relation. As a result of the adaptation to oppositely rotated prisms, retinal directions give rise to altered subjective directions and thus effect a change in the set of corresponding retinal points.[25] Ogle, however, appears to believe that the set of retinally corresponding points is a "preformed and stable arrangement." Given the outcome of this and other experiments, such a view seems untenable. The set of retinally corresponding points appears modifiable.

---

[24] J. Hochberg, *Perception*, 1964.

[25] Adaptation to monocular visual-field rotation appears to involve a change in the association between retinal and egocentric direction; see A. Mack and I. Rock, A re-examination of the Stratton effect: Egocentric adaptation to a rotated visual image, *Percept. & Psychophys.*, 4, 1968, 57–62.

The fact that there was no evidence of cyclotorsion in observers who showed adaptation to the imposed distortions of orientation would seem to rule out an oculomotor account for at least this kind of perceptual adaptation. In order for cyclotorsion to have accounted for the adaptive changes in the perception of upright (an adaptation which seems most appropriate to cyclotorsional explanation), the eyes would have to have rotated .91° in the opposite directions. In fact, the recorded mean combined cyclotorsion (CLRC) was .004°.

Because the absence of any evidence of cyclotorsional eye movements seemed surprising, cyclotorsion was measured in normal (nonadapted) observers while they inspected a line inclined 15° backward, an objectively upright line, a line subjectively upright, and a line inclined forward by 6°. Analysis of the photographs taken during these observations revealed no evidence of disjunctive cyclotorsional eye movements during inspection of an inclined line. It therefore appears that inspection of an inclined line was not a sufficient condition for disjunctive cyclotorsion. This finding, while in itself surprising, makes the data indicating no cyclotorsion in experimental observers less surprising. Our findings are in full agreement with those reported by Howard and Templeton, who after photographing the eyes of observers viewing a line inclined by 10° found that "not one of the three [observers] showed any evidence of a significant mean change of the torsional position of the eye that could be related to the angular position of the stimulus line."[26]

In sum, because cyclotorsion did not occur and because we failed to get evidence of adaptation to either two- or three-dimensional distortions of orientation in observers who were not permitted to move about during binocular prism exposure, we conclude that eye-dependent perceptual learning did occur and that exposure to oppositely rotated visual fields led to adaptive changes in the relation between retinal disparity and perceived depth.

---

[26] I. P. Howard and W. B. Templeton, Visually induced eye torsion and tilt adaptation, *Vis. Res.*, 4, 1964, 433–437.

THE EFFECT OF SEPARATION BETWEEN TEST AND COMPARISON OBJECTS ON SIZE CONSTANCY AT VARIOUS AGE-LEVELS

RICHARD W. BRISLIN and H. W. LEIBOWITZ,
Pennsylvania State University

Abstract. The effect of separation between test and comparison stimuli on size constancy at various distances was determined for children and adults. Separation slightly reduced the tendency toward size constancy at the longer distances. The data are interpreted as supporting the assumption that size constancy and the Ponzo illusion share common mechanisms, since the developmental size-distance function is independent of whether the comparison is successive, as in most size-constancy experiments, or simultaneous, as in the Ponzo illusion.

A number of investigators have suggested that size constancy and the Ponzo perspective illusion share common mechanisms.[1] Specifically, it is assumed that one of the factors underlying size constancy is that subjects make depth inferences from converging lines, which are ordinarily associated with distance in the normal environment. When converging lines are presented in the two-dimensional Ponzo figure, distance is falsely indicated, resulting in an inappropriate correction of the upper or 'more distant' line of the Ponzo figure. Evidence for this hypothesis is presented by the similar increase in both measures as a function of chronological age.[2]

Received for publication December 2, 1969. The research was supported by Grant MH-08061 from the National Institute of Mental Health. Dr. Brislin is now at Western Washington State College.

[1] R. Tausch, Optische Täuschungen als artifizielle Effeckte der Gestaltungsprozesse von Grössen- und Formenkonstanz in der naturlichen Raumwahrnehmung. *Psychol. Forsch.*, 24, 1954, 299–348; E. von Holst, Aktive Leistungen der menschlichen Gesichtswahrnehmung, *Stud. Gen.*, 10, 1957. R. L. Gregory, *Eye and Brain*, 1966, 145–156.

[2] H. W. Leibowitz and J. M. Judisch, The relation between age and the magnitude of the Ponzo illusion, this JOURNAL, 80, 1967, 105–109; H. W. Leibowitz, S. W. Pollard, and D. Dickson, Monocular and binocular size matching as a function of distance at various age levels, this JOURNAL, 80, 1967, 263–268.

However, one criticism which might be leveled at this argument is that when one views the Ponzo figure, both elements to be compared are visible simultaneously, while in almost all size-contrary experiments, including those referred to above, the test and comparison objects are separated by 90° and are therefore necessarily viewed successively. In order to determine whether viewing angle is in fact a variable in the developmental function of size constancy, the present study was planned to investigate both simultaneous and successive comparison. This was accomplished by positioning the test and comparison stimuli at an angular separation of 5.75° (i.e., simultaneous comparison), in addition to the more conventional 90° separation.

METHOD

Stimuli. The test objects consisted of white disks 3.0, 6.0, 12.0 and 24.0 in. in diameter, which were viewed at 25, 50, 100, and 200 ft., respectively, from *O*. Thus, the visual angle subtended by each test object was constant at .57°. The test objects were mounted on vertical wooden supports constructed so that not only the disk but the frame and base were proportional to viewing distance. The disks were painted flat white (reflectance, 73%), while supports and bases were painted flat black (reflectance, 4.9%). (Details of the construction of the stimuli have been previously described.)[3] The comparison stimuli consisted of a series of circles cut from white cardboard (reflectance, 72.2%) and ranging in size from .5-in. to 33.0-in. diameters (in ½-in. steps up to 7 in. and in 1-in. steps thereafter). The circles were viewed against a 4 × 8 ft. vertical background covered with black flock paper (reflectance, .85%). The stimuli were attached to the center of this background by means of 'burr' tape. The centers of the test as well as the comparison stimuli were 48 in. from the ground.

Data were obtained under two conditions of angular separation between the test and comparison stimuli. In the more conventional arrangement, which served as the control condition for the present study, these were located 90° from each other and the *O* was required to make head movements in order to alternately view the test and comparison stimuli. In the experimental condition, the angle between test and comparison stimuli was made as small as possible. Because of the width of the background against which the comparison stimuli were viewed, this minimum angle was 5.75°. The comparison stimuli were located 25 ft. from the *O*. Each *O* viewed the test stimuli at all four distances.

Procedure. The testing was carried out in a local playground on warm, sunny days between 1:30 and 3:30 P.M. DST, with the sun at the *O*'s back. The *O*s were 36 children (ranging in age from 5 to 12 years) who frequented the playground, and 12 university students. Of these, 18 children and 6 students formed four age-groups of six subjects each to provide the data for the experimental condition. (5.75° separation). The remaining 18 children and 6 college students made up four groups for the control condition (90° separation). Assignment to experimental or control groups was on a random basis. All *O*s were free from visual defects, and there were equal numbers of males and females in each age group. The mean ages of the four 5.75° groups matched the mean ages of the four 90° groups.

---

[3] H. W. Leibowitz and R. A. Dato, Visual size-constancy as a function of distance for temporarily and permanently monocular observers, this JOURNAL, 79, 1966, 279-284.

The children were asked if they wanted to 'play a game,' by a female assistant. The *O* was then escorted from the main playground area to the testing area out of view of the other children. The task as explained to the *O* was simply to indicate which of the two circles was 'bigger.' No time limit was placed on the *O*, and the test was completed in approximately 20 min.

In order to make sure that the *O* understood the instructions, the first comparison stimulus presented was clearly larger or smaller than the test stimulus. This was followed by another comparison chosen to elicit the opposite response. This procedure was repeated several times until it was clear that the *O* understood the instructions, after which the difference between stimuli was systematically decreased. With this procedure it was possible to determine the size of the comparison stimulus at which the transition from smaller to larger occurred and from which thresholds were calculated. The response 'equal' was allowed when offered. The matching value was taken as the midpoint of the uncertainty interval of the *O*'s responses. It was not necessary to present all possible comparison stimuli to the *O* on a given trial, since typically the transition was between successive stimuli in the series.

RESULTS

The results are plotted in Figure 1 as matched size as a function of test-object distance for the various age-groups and the two separation conditions. For both separation conditions, the pattern observed in previous studies was replicated; that is, the various age-groups made similar matches at the short observation distance, but with increasing distance the children showed less tendency toward constancy than did the adults. The younger the children, the more their matches deviated from size constancy at the longer observation distances.

Of particular interest is the comparison between the two separation conditions. It can be observed that for the groups of children (6.0, 9.4, and 12.5 years) the matches for the 90° separation were generally lower than for the corresponding age-groups with the 5.75° separation, but only at the longer distances.

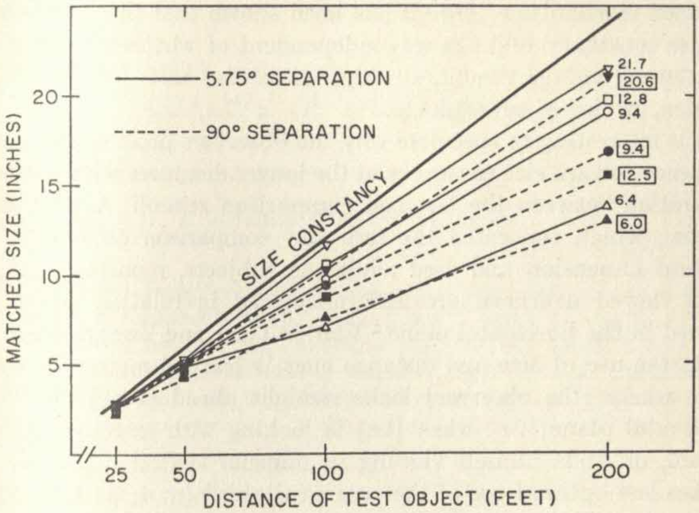


FIG. 1. Matched size as a function of distance at various age-levels for two separations between test and comparison objects

Analysis of variance revealed that the tendency toward constancy increased with age ($F = 8.5$, $df = 3, 40$, $p < .01$). The distance of the test object was highly significant ($F = 675.5$, $df = 3, 120$, $p < .001$), and the interaction of age and distance was also significant ($F = 8{,}6$, $df = 9, 120$, $p < .01$). The main effect of the two separations reached only the .1 level ($F = 3.8$, $df = 1, 40$), while the interaction of separation and distance reached the .05 level ($F = 2.8$, $df = 3, 120$). Other interaction effects were not significant.

DISCUSSION

The principal result of the present study is that the tendency toward size constancy increases as a function of age and independent of the angle of separation. This increase, however, was found to be slightly greater at the longer distances for subjects who observed at the 5.75° separation than for those who observed with the more conventional 90° separation between test and comparison objects. Thus, the previous finding that size constancy improves with age for distant but not for near objects was here replicated for two angles of separation. These data also are consistent with the reasoning that the Ponzo illusion and size constancy share common mechanisms, since it has been shown that the relationship of size constancy and age was independent of whether the test objects were viewed simultaneously, as is the case for the Ponzo illusion, or viewed successively.

It is interesting to speculate why the observers produced a greater tendency toward size constancy at the longer distances with a smaller separation between the test and comparison stimuli. A number of studies, which separated the test and comparison objects in the *vertical* dimension and used adults as subjects, reported that objects viewed overhead are underestimated in relation to objects viewed in the horizontal plane.[4] Van de Geer and Zwaan suggested that "the use of size and distance cues is learned mainly in situations where [the observer] looks straight ahead at objects in the horizontal plane; . . . when [he] is looking with extreme angle of regard, or finds himself viewing in unusual spatial directions, he makes less optimal use of the cues available" (n. 4; p. 571). Data supporting this hypothesis have been provided by Leibowitz and Hartman for children's observations in the vertical plane.[5] In line with this interpretation, it would be expected that since one has generally more experience comparing distant objects viewed straight ahead than at a 90° separation, matches made between spatially adjacent objects should be more accurate than when the objects to be compared are separated angularly. The general principle suggested by these results is that separation of test and comparison objects produces a diminution of the tendency to size constancy in situations in which the cues are poor or have not been completely learned by the observer.

[4] E. Schur, Mondtäuschung and Sehgrossenkonstanz, *Psychol. Forsch.*, 7, 1926, 44–81; J. P. Van de Geer and E. J. Zwann, Size-constancy as dependent upon angle of regard and spatial direction of the stimulus-object, this JOURNAL, 77, 1964, 563–575; Ching Chi-cheng, Peng Jui-hseung, Fang Yun-chiu, and Lin Chung-hsien, *Acta Psychol. Sinica,* 24, 1963, 175–185.

[5] H. Leibowitz and T. Hartman, Magnitude of the moon illusion as a function of the age of the observer, *Science,* 130, 1959, 569–570.

COVERT ORAL BEHAVIOR AS A FUNCTION OF QUALITY OF HANDWRITING

F. J. McGuigan, *Hollins College*

Abstract. Two experiments were conducted in which subjects wrote words or engaged in comparable nonlinguistic (control) tasks. It was found that both covert oral behavior and breathing rate, but not covert nonoral behavior, were significantly greater during the linguistic task. Furthermore, amplitude of covert oral behavior varied inversely with the quality of handwriting. The interpretation is that covert oral behavior is beneficial in the performance of language tasks.

There is considerable evidence that subjects make covert oral responses when engaged in a variety of language tasks. For instance, heightened levels of electromyograms (EMG) have been recorded from several oral regions (tongue, lips, or chin) when subjects silently read prose, when they listened to prose, when they solved problems, and when they thought about abstract terms such as 'electrical resistance.'[1]

Handwriting is a linguistic activity for which associated covert behavior has been less extensively studied. The implication from the performance on other linguistic tasks is that subjects would

---

Received for publication September 5, 1968. Edited by Professor E. B. Newman. The research reported herein was performed pursuant to a contract with the United States Office of Education, Department of Health, Education, and Welfare, under the provisions of the Cooperative Research Program. The author thanks Douglas Gresham, Susan Guthrie, Marty Muller, Norman Ostrov, and Ronald Savukas for their help in collecting and analyzing data.

[1] A. W. Edfeldt, *Silent Speech and Silent Reading,* 1960, 153; F. J. McGuigan, B. Keller, and E. Stanton, Covert language responses during silent reading, *J. educ. Psychol.,* **55**, 1964, 339–343; F. J. McGuigan and W. I. Rodier III, Effects of auditory stimulation on covert oral behavior during silent reading, *J. exp. Psychol.,* **76**, 1968, 649–655; A. A. Smith, R. B. Malmo, and C. Shagass, An electromyographic study of listening and talking, *Canad. J. Psychol.,* 8, 1954, 219–227; H. Wallerstein, An electromyographic study of attentive listening, *Canad. J. Psychol.,* 8, 1954 228–238; L. A. Novikova, Electrophysiological investigations of speech, in N. O'Connor (ed.), *Recent Soviet Psychology,* 1961, 210–226; E. Jacobson, Electrophysiology of mental activities, this JOURNAL, 44, 1932, 677–694.

make covert oral responses during cursive writing, too. Furthermore, since it has been found that poor readers have a higher level of covert oral behavior during silent reading than do good readers,[2] it could be supposed that covert oral behavior during handwriting would be greater for poor writers than for excellent writers. This is, in fact, precisely the conclusion that Lepley reached, though Lepley did not use direct measures of covert behavior during writing.[3] Rather, he classified his subjects according to whether they said that they subvocalized when writing words. He found that those who said they did subvocalize had a poorer quality of writing than did those who said they did not.

Lepley's interpretation of his findings was that when "implicit speech accompanies cursive writing, the motor rhythm and harmony thereof are disturbed," though he recognized that the influence may be in the reverse direction—"that individuals deficient in motor coordination may attempt to support and guide their acts of skill with symbolic action" (p. 599). Lepley's quandary is typical of those who are concerned with covert speech during the performance of language tasks, and there is considerable controversy about the function of such behavior.[4] In addition to the possibilities that covert oral responses are detrimental to the performance of a linguistic task, or that they facilitate the reception and processing of linguistic stimuli, it is also possible that they are merely one aspect of a heightened state of general arousal. It is, therefore, important to compare covert activity during cursive writing with such activity during a comparable motor task that does not involve language, such as the drawing of ovals. Furthermore, relative changes in oral and nonoral behavior during the performance of a language task should be studied in order to ascertain whether it is the speech mechanism per se that is involved. If covert oral behavior is greater during the performance of a linguistic task than during a comparable nonlinguistic task, and if increases in covert behavior during the linguistic task are localized in the speech region, then it seems likely that those covert oral responses serve a linguistic function.

It was the purpose of this research to test the hypotheses that

---

[2] Edfeldt (n. 1).

[3] W. M. Lepley, The participation of implicit speech in acts of writing, this JOURNAL, 65, 1952, 597–599.

[4] F. J. McGuigan, *Thinking: Studies of Covert Language Processes*, 1966, 293.

(*a*) subjects increase their covert oral behavior during cursive writing; (*b*) amplitude of covert oral behavior during cursive writing varies inversely with the quality of handwriting; (*c*) covert oral behavior increases more during writing than during the performance of a comparable nonlinguistic task; and (*d*) changes in covert behavior during writing are localized primarily in the oral region. The results of two experiments—and one minor control study, see n. 11—are reported.

EXPERIMENT I

Method

Subjects. The handwriting of 117 women in introductory psychology classes at Hollins College was rated by 30 judges. Seven *S*s who had excellent handwriting were assigned to Group E, and seven who had poor handwriting to Group P. Inability to cooperate later eliminated three *S*s from Group P.

Procedure. Lepley's procedure (n. 3) was used except as noted below. The *S*s were brought individually to the laboratory and the following sensors were attached: Grass electrodes were attached to the left (nonwriting) arm and to the chin;[5] specially constructed vacuum electrodes were placed on the tongue, and a bellows pneumograph about the chest. Each *S* then assumed a relaxed, comfortable writing position at a table and was told always to keep her eyes on a reading stand in front of her. The *S*s briefly practiced writing words and ovals prior to the experimental session. There followed three experimental periods of 1 min. each; during the first, *S* drew ovals; during the second, wrote words; and during the third, drew ovals again.[6] Cards containing words or ovals to be copied were exposed (as appropriate) on the reading stand by removing a blank card before each period. A rest period of 1 min. preceded each experimental period, and *S*s were instructed immediately prior to each period what they were to write. So that *S* would not move her arm, she wrote the words or ovals directly over each other during each period; the length of hand movement was the same for both conditions. The 10 words used by Lepley were supplemented by an additional 18 of the same character for the period of writing.

Apparatus. The laboratory consisted of two adjacent sound-deadened but electrically unshielded rooms, one for the *S*, and one for observation and for the recording apparatus. Each sensor led into two Tektronix 122 amplifiers and a Honeywell Galvanometer amplifier, all placed in series to provide an amplification of 10,000. Signals were recorded on a Honeywell Visicorder, and the EMG measures were also recorded on three coordinated audio tape recorders.

---

[5] J. F. Davis, Manual of surface electromyography, WADC Technical Report 59-184, Wright Air Development Center, Ohio, December 1959.

[6] This design was used in Experiment I to explore the relative effects of drawing ovals before and after writing words by the same *S*s. Essentially the same results were obtained regardless of whether ovals preceded or followed the word period. For brevity, only data for the second oval period are reported.

Quantification of the data. The analog EMG signals from the magnetic tapes entered a root mean-square (RMS) voltmeter, which produced a d.c. signal proportional to the true RMS value of the input signal. The output of the RMS voltmeter was read instantaneously every 5 sec. by a digital voltmeter and printed out on a digital recorder. These values represent, then, root mean-square voltages at 5-sec. intervals. Values obtained 5 sec. before and after each set of instructions were discarded. The mean values were computed for each *S* for each period and each rest session. Rate of respiration was quantified as number per minute.[7]

Results

We first sought to determine whether the subjects, regardless of their group, showed greater covert behavior during writing than during the (prewriting) rest session; this was accomplished separately for each measure (tongue, chin, arm EMG, and respiration rate). For this purpose, mean response values during resting were subtracted from mean response values during writing for each subject. A group mean of these differences was then determined for each measure. Similar group mean differences were obtained by subtracting values during rest from the corresponding values during the oval periods. The resulting group means are presented in Table I.[8] It can be seen that the amplitudes of the tongue and chin responses significantly increased during writing, but these responses did not significantly increase while drawing ovals. For example, the mean difference between rest and writing for tongue EMG was 9.5 μv. ($p < .05$; α was set at .05 throughout), but for drawing ovals this measure increased by only 5.3 μv. ($p > .05$). The mean EMG of the nonwriting arm did not significantly increase during writing or during the drawing of ovals. Breathing rate, on the other hand, significantly increased during both periods.

---

[7] For some details of these procedures, see the method used in the second experiment of McGuigan and Rodier (n. 1).

[8] These mean differences were tested to determine whether they were significantly different from zero by means of Sandler's *A* test. The *A* test leads to precisely the same conclusions as the paired *t* test but is preferable to it for several reasons; see F. J. McGuigan, *Experimental Psychology: A Methodological Approach* (2d. ed.), 1968, 176–178.

TABLE I

MEAN RESPONSE INCREASES DURING EXPERIMENTAL PERIODS RELATIVE TO REST SESSION: EXPERIMENT I

	Writing − Rest		Ovals − Rest		Difference	
Measure	$\bar{X}$	A	$\bar{X}$	A	$\bar{X}$	A
Tongue EMG (μv.)	9.5	.147*	5.3	.313	4.2	.270*
Chin EMG (μv.)	3.9	.193*	.4	3.537	3.5	.194*
Arm EMG (μv.)	2.7	.965	1.1	1.29	1.6	1.600
Respiration (per min.)	2.7	.147*	1.4	.234*	1.3	.388

Note: $N = 11$.
* $p < .05$.

To determine whether the increase in each response measure was greater during handwriting than while drawing ovals, each mean increase under the oval condition was subtracted from the mean increase while writing for each subject; the resulting means appear in the difference column of Table I. For instance, the mean increase in chin EMG during writing was 3.9 μv., during the oval period it was .4 μv.; and the difference between these two values is 3.5 μv. ($p < .05$). These results indicate that subjects increased the amplitude of their covert oral responses (tongue and chin EMG) significantly more during the writing than during the oval period. On the other hand, the measure of covert nonoral behavior (EMG from the inactive arm) did not increase significantly more during handwriting than while drawing ovals; nor was there a significant difference in breathing-rate change as a function of experimental condition.

To test the hypothesis that amplitude of covert oral behavior varies inversely with the quality of a person's handwriting, the results in the difference column of Table I were analyzed according to whether the subject's handwriting was judged to be excellent or poor. We can see in Table II that the means for the two measures of oral behavior (tongue and chin EMG) were higher for Group P (poor writers) than for the excellent writers (Group E); but t tests between groups show that the differences were not significant ($t = .72$ and $.62$ for the tongue and chin measures respectively). The differences in the arm and respiration measures between Group E and Group P were minor and did not approach significance.

TABLE II

Mean Difference between Response Increases While Writing and Drawing Ovals: Experiment I

Measure	Group E		Group P	
	$\bar{X}$	*A*	$\bar{X}$	*A*
Tongue EMG (μv.)	3.1	.559	6.2	.521
Chin EMG (μv.)	2.8	.222*	4.7	.565
Arm EMG (μv.)	1.9	2.48	1.1	3.08
Respiration (per min.)	1.3	.656	1.2	1.320

* $p < .05$.

EXPERIMENT II

Method

Subjects. Groups P and E contained eight *S*s each, selected from a larger group of 90 college students (women) as in Experiment I. Two additional (overlapping) groups of eight *S*s each were also formed such that one group consisted of *S*s who said that they subvocalized and one group of *S*s who said that they did not.[9]

Procedure. The order of the writing and the oval periods was counterbalanced for each group, so that half of each group wrote words before they drew ovals, while the remaining half drew ovals before they wrote words. The same measures of covert responses were recorded as for Experiment I except that leg EMG was substituted for arm EMG and that electroencephalographic (EEG) measures from the right motor area (C-4) were taken. All other procedures of Experiment I were employed, including rest sessions prior to each experimental period.

Apparatus. The laboratory consisted of subject rooms and a central apparatus (control) room, all with built-in high-permeability steel and magnetic liners to effectively shield radio and low-frequency signals. The amplifiers were the same as for Experiment I, but all signals were recorded on an eight-track data tape recorder.

Quantification of the data. The analog EMG and EEG signals from the tape recorder entered an RMS voltmeter, as before; the resulting signals entered a voltage-to-frequency converter which could vary between 0 and 10,000 counts/sec, and thence to an electronic counter which counted the frequency for each 10-sec. period. The resulting frequency was converted to a binary-coded decimal signal that was printed out on a digital recorder; this value is a mean integrated voltage and provided the amplitude of each measure for each 10-sec. interval of the experiment. The values of the first and last 10 sec. of each experimental period were discarded, and means for each period were computed using the remaining values. At the conclusion of the experiment, each *S* was requested to make overt responses involving each measure being recorded; the maximum value for a 10-sec. interval during the experiment was ascertained, this usually when *S* intentionally made the overt response. Similarly, the minimum value for a 10-sec. period was determined for each *S*. These data were then used to compute a ratio for each *S* as follows:[10]

---

[9] These groups were formed on the following basis. Two *S*s from Group E and two from Group P said that they subvocalized during cursive writing; four *S*s from Group E and two from Group P said that they did not. The remainder classified themselves as 'doubtful.' Consequently, four additional *S*s who said that they subvocalized and two who said that they did not were randomly selected from the larger group of 90 *S*s in order to form the two additional groups of eight *S*s each who said that they did or did not subvocalize.

$$\text{Ratio} = \frac{\bar{X} - \text{minimum}}{\text{Maximum} - \text{minimum}}.$$

For example, the mean tongue amplitude ($\bar{X}$) was computed for the writing period for an *S*, and the minimum tongue amplitude during the entire experiment for that *S* was subtracted from that mean; the resulting value was then divided by that *S*'s maximum value for a 10-sec. interval subtracted from that *S*'s minimum value. The result is a ratio that can vary between 0 and 1.0 such that the higher the value, the greater *S*'s response amplitude. Lykken et al. (n. 10) have shown that this ratio corrects for individual differences in *S*'s range of response values, and thus yields a more meaningful dependent variable value than that typically used.

In short, a mean ratio was computed for each EMG and EEG measure for the writing and for the oval periods for each *S*. Pneumograms were quantified by determining respirations per minute during the resting sessions and during the experimental periods for each *S*.

Results

Group mean ratios based on all subjects were computed for each experimental period for the EMG and EEG measures and entered in Table III. We can note, for example, that the mean ratio of tongue EMG for all subjects was .439 while writing words and that it was .293 while drawing ovals. The difference between these two means (.146) was tested to determine whether the value during writing was significantly different from the value during the oval period. As can be observed, the subjects yielded a significantly

---

[10] This ratio is a simplified statement of that reported by D. T. Lykken, R. Rose, B. Luther, and M. Maley, Correcting psychophysiological measures for individual differences in range, *Psychol. Bull.*, 66, 1966, 481–484.

TABLE III

Means during Experimental Periods: Experiment II

Measure	Words	Ovals	Difference	*A*
Tongue EMG	.439	.293	.146	.101*
Chin EMG	.430	.362	.068	.558
Leg EMG	.377	.345	.032	3.440
EEG	.397	.385	.012	23.500
Respiration	3.82	2.60	1.22	.162*

Note: Respiration is rate of increase relative to rest session; others are ratios. $N = 22$.

* $p < .05$.

higher amplitude of tongue response when writing words than when drawing ovals, thus confirming the results of Experiment I. While chin EMG was noticeably higher during the word period, this difference was not significant. Neither difference in the nonoral measures (leg EMG or EEG) was sizable or significant. Respiration rate, on the other hand, increased significantly more during the writing period than it did during the oval period. The differences between rest and experimental periods for mean integrated EEG were minor and not significant (for words the mean increase was .10 μv. and for ovals it was .25 μv.).

To study the results as a function of quality of handwriting, the difference scores in Table III were classified according to whether the subject was a member of Group P or E. That is, the ratio for each subject while drawing ovals was subtracted from her ratio while writing words, and means of these differences for Groups P and E were computed. It can be observed in Table IV that the mean

TABLE IV

Differences between Means While Writing and Drawing Ovals: Experiment II

Measure	Group P	Group E	Difference	*t*
Tongue EMG	.243	.061	.182	2.36*
Chin EMG	.148	.002	.146	1.33
Leg EMG	.061	.009	.052	.38
EEG	−.030	.110	.140	.78
Respiration	2.19	1.38	.82	.76

Note: $N = 8$ per group.

* $p < .05$.

tongue EMG for Group P was .243 higher while writing words than while drawing ovals; the difference for Group E was only .061. The difference between these two differences (.182) was significant, allowing us to conclude that the poor writers yielded a greater amplitude of tongue EMG during writing (relative to drawing ovals) than did the excellent writers. Group P's increase in chin EMG was also noticeably higher than was Group E's, but the difference failed to be significant. There were only minor and nonsignificant differences between groups for leg EMG, EEG, and respiration rate. It may be concluded that poor writers emitted a significantly greater amplitude of covert oral behavior when writing words than did excellent writers, but they did not differ significantly as far as covert nonoral behavior, EEG, or respiration rate are concerned.

The results presented in Table V were computed just as for Table IV, except that the classification is by whether the subjects said that they do or do not subvocalize. The question here was whether individuals who say that they subvocalize during cursive writing emit a larger amplitude of covert oral behavior during writing than do subjects who deny that they subvocalize. No behavioral differences between these groups approached significance.

TABLE V

DIFFERENCES BETWEEN MEANS WHILE WRITING AND DRAWING OVALS: EXPERIMENT II

	Subvocalizers			
Measure	'Yes'	'No'	Difference	t
Tongue EMG	.170	.099	.071	1.01
Chin EMG	.036	.124	−.088	.72
Leg EMG	−.031	.141	−.172	1.30
EEG	−.060	.180	.240	1.85
Respiration	.84	1.12	−.28	.25

Note: $N = 8$ per group.

DISCUSSION

The major findings of Experiments I and II may be summarized as follows. With regard to covert oral behavior (tongue and chin EMG), it may be concluded that (*a*) the subjects significantly increased the amplitude of their covert oral responses during writing but not during a control (nonlinguistic) task; (*b*) the amplitude of covert oral behavior during writing was significantly greater than during the performance of the nonlanguage task; and (*c*) amplitude of covert oral behavior during writing varied inversely with the quality of a person's handwriting. With regard to covert nonoral behavior (arm and leg EMG), it may be concluded that (*d*) the subjects did not significantly increase the amplitude of their responses during writing or during the drawing of ovals, nor was there a significant difference in behavior as a function of these tasks.

These findings suggest the following conclusions. First, the fact that increased covert behavior during writing occurred in the oral but not in the nonoral regions sampled suggests that these responses were not widely distributed throughout the body; therefore the heightened covert oral behavior was apparently not merely an aspect of general arousal of the entire body but was relatively localized in the speech mechanism. Second, the findings that covert oral behavior was significantly greater in amplitude during writing than during the performance of a comparable nonlinguistic task indicates that those responses were intimately associated with the performance of a linguistic task;[11] the question is, in what manner?

---

[11] The problem of selecting a nonlanguage task comparable to the writing of words as far as bodily arousal, movements, etc., are concerned is a difficult one. As will be seen, the fact that tongue EMG was significantly higher while writing words than while drawing ovals suggests that the heightened covert oral behavior served a language function in that it was involved in the reception and processing of language stimuli. Nevertheless, it could be argued that the relatively heightened tongue EMG occurred, for example, merely because writing words requires more variations in the direction of hand movements than does the drawing of ovals. The question is, what other nonlanguage task might be more similar to writing words, but not so novel as to produce considerable covert behavior merely because it requires more effort or attention than copying familiar words? To study this problem, a third, control, experiment was conducted with six subjects randomly selected from psychology classes. The same procedures as reported above were used, except that the subjects wrote words, drew ovals, or drew clefs, in counterbalanced order. Unfortunately, the subjects were not as adept at drawing clefs as they were at the other tasks, so that drawing clefs is probably too demanding to serve as a suitable control task. Nevertheless, and even though the *N*s were too small to allow us to run meaningful tests of significance, the results were in line with those of Experiments I and II: the mean ratio for tongue EMG while writing words was .737, while for drawing clefs and ovals it was .603 and .564 respectively. These findings, taken with

Two possible functions of heightened covert oral behavior during the performance of a language task have been stated: that such responses are detrimental to a person in that they interfere with his language proficiency; and that the heightened covert oral behavior is beneficial in that it facilitates the reception, decoding, or processing of incoming language stimuli. The former is the classical position and, with regard to subvocalization during silent reading, has been stated in such ways as: "Any observable form of vocalization . . . such as silent lip movement . . . retards the rate of silent reading [This] has been common professional knowledge since the early scientific studies of reading."[12] Lepley, it will be recalled, interpreted his findings in this manner. With regard to the present study the issue is: did Group P exhibit a poorer quality of handwriting than Group E because their relatively greater amplitude of covert oral behavior interfered with writing proficiency? Or do poor writers find writing sufficiently more exacting that they need to exaggerate their level of covert oral behavior in order to more adequately perform? The fact that subjects generally tend to increase their level of oral responding during the performance of linguistic tasks argues that such responses are beneficial. If they are beneficial, then it would follow that the more difficult the task, the greater the need to exaggerate covert oral responses. If this hypothesis is tenable, it should be possible to manipulate experimental conditions and observe systematic variation of the amplitude of covert oral behavior. Several experiments have been successfully conducted with this as their purpose. For example, Edfeldt showed that amplitude of silent speech increases as the difficulty of the prose being read is systematically increased. Edfeldt's interpretation is that silent speech occurs in all individuals and that it is likely that "silent speech actually constitutes an aid toward better reading comprehension" (n. 1; p. 154). A similar result is reported by Sokolov, though in the translation cited he provides data for but a single case; the conclusion is that amplitude of lip EMG increases as the subject moves from writing familiar material to unfamiliar (and more difficult) material.[13] One further example is provided

---

the fact that the difference in tongue EMG between writing words and drawing ovals was higher for Group P than for Group E, strengthen the interpretation that the covert oral behavior served a language function.

[12] E. A. Betts, *Foundations of Reading Instruction,* 1950, 450.

[13] A. N. Sokolov, Speech-motor afferentation and the problem of brain mechanisms of thought, *Soviet Psychol.,* 6, 1967, 10.

by the work of McGuigan and Rodier (n. 1). These experimenters demonstrated that amplitude of covert oral behavior increased when several kinds of auditory stimulation were introduced into the silent reading situation. Their interpretation was that the increased amplitude of covert oral behavior facilitated silent reading.

In short, these findings indicate that covert oral behavior is beneficial in the performance of linguistic tasks; it remains to more precisely determine the role of this class of behavior in the reception, decoding, or processing of language stimuli.

Previous research has shown that respiration rate increases significantly during silent reading.[14] We have noted here that this measure increased significantly during both the oval and writing periods, and that the increase was significantly greater while writing than while drawing ovals. Since respiration is intimately involved in the production of speech, it may be that the relatively fast breathing rate during linguistic tasks occurs because subjects make covert *language* responses—because they are in fact engaged in 'silent speech' during writing and reading.

---

[14] McGuigan, Keller, and Stanton (n. 1); McGuigan and Rodier (n. 1).

THE EFFECT OF ACOUSTIC CONFUSABILITY ON VISUAL SEARCH

LESTER E. KRUEGER, *Harvard University*

Abstract. Subjects searched more slowly through same-sounding visual displays in which the letters, when pronounced, matched the vowel sound of the target letter (e.g., a target of C and a display of B D C V B P) than through different-sounding displays (e.g., a target of C and a display of F X C S L M), indicating that subjects do form acoustic representations of display letters in visual search. Subjects also made more errors on same-sounding than different-sounding displays. Search was slowed to a greater extent on same-sounding (versus different-sounding) displays for subjects who more frequently subvocalized the target letters.

The present experiments investigate how deeply subjects encode letters in a visual display while searching for a predesignated target letter. Specifically, they investigate whether subjects form acoustic representations of the display letters. Recent evidence prompted Sperling to postulate that visual stimuli are transferred at an early stage from a rapid-decay visual information store to a slower-decay auditory information store.[1] Conrad, for example, found the same types of letter confusions in written recall following visual presentation as he found in a listening task.[2] For instance, in both tasks letters whose pronunciation ends in ē (e.g., B, C, D) tended to be confused with each other, as were letters whose pronunciation begins with ĕ (e.g., F, X, S). Wickelgren presented items auditorily and similarly found that intrusions in recall among letters and between letters and numbers tended to follow acoustic similarity.[3]

Received for publication March 4, 1969. Edited by Professor E. B. Newman. This study was supported by a predoctoral traineeship by the United States Public Health Service, National Institute of General Medical Sciences, 5-T01-GM01011-07, to Harvard University. The author is now at the City College of City University of New York.

[1] G. A. Sperling, A model for visual memory tasks, *Hum. Factors*, 5, 1963, 19–31.

[2] R. Conrad, Acoustic confusions in immediate memory, *Brit. J. Psychol.*, 55, 1964, 75–84.

[3] W. A. Wickelgren, Acoustic similarity and intrusion errors in short-term memory, *J. exp. Psychol.*, 70, 1965, 102–108.

However, when the task is changed from recall to one involving the immediate report of tachistoscopically presented letters, acoustic confusability no longer seems to affect performance. Glucksberg, Fisher, and Monty presented very briefly three rows of six letters each, and immediately before or after the presentation, told the subject which row to report.[4] Letters in a particular row were all acoustically similar (e.g., F, S, X) or dissimilar (e.g., J, K, N). They also were all visually similar or dissimilar. High visual similarity, or confusability, led to greater errors in the subjects' reports, but high acoustic confusability did not. Glucksberg, Monty, and Fisher used much the same experimental procedure and found significantly superior, rather than inferior, performance on rows of acoustically similar letters. They also found a significant interaction between acoustic confusability and visual confusability, "suggesting that the acoustic confusability effect may be attributable to the particular selection of visually confusable items."[5]

In visual search, which is the focus of the present investigation and in which the demands on immediate memory are minimal, much evidence indicates that subjects do not encode display letters so deeply as to form acoustic representations. Neisser's well-practiced subjects could search as fast for ten target letters as for a single letter, and they reported afterward that they did not 'see' the nontarget letters.[6] Neisser concluded that his subjects processed only enough features to eliminate display letters from consideration and that nothing may take place "at the level of letter identification until the target letter appears."[7] However, Neisser pressed his subjects to search ever faster, at the cost of a high error rate; if they had been cautioned to be accurate as well as fast, they might have encoded letters more fully.

Sternberg eliminated error-prone subjects, yet his evidence, too, indicates that display characters are represented in terms of "physical properties of the stimulus, rather than the identity or name of the character."[8] Sternberg found that superimposing a checkerboard grid on display digits slowed their comparison with the memory image of the target letters, which indicates there was residual degradation in the stimulus representations of the display digits. However, stimulus degradation impeded encoding even more than memory comparison; and with normal, intact display digits the subjects may yet form representations in terms of the name or sound of the character.

---

[4] S. Glucksberg, D. F. Fisher, and R. A. Monty, Brief visual memory as a function of visual and acoustic confusability, *Proc. Amer. psychol. Assoc.*, 2, 1967, 55–56.

[5] S. Glucksberg, R. A. Monty, and D. F. Fisher, Brief visual memory: Interference or decay? unpublished paper, n.d.

[6] U. Neisser, Decision-time without reaction time: Experiments in visual scanning, this JOURNAL, 76, 1963, 376–385; U. Neisser, R. Novick, and R. Lazar, Searching for ten targets simultaneously, *Percept. mot. Skills*, 17, 1963, 955–961; U. Neisser, Visual search, *Sci. Amer.*, 210, 1964, 94–102.

[7] Neisser (n. 6; *Sci. Amer.*), 98.

Whether subjects form acoustic representations of display letters during visual search can be directly tested by determining whether subjects search slower through lists of letters which, when pronounced, sound like the target letter, than through lists of letters which have different sounds. That is, if the subjects form acoustic representations, then display letters which sound more similar to the target letter would be examined more fully, and longer, before being discarded. Increased similarity of stimuli increases reaction time in card-sorting tasks,[9] and Neisser and Beller found slower search when the display letters had more geometrical features in common with the target letter.[10]

Several studies found no effect on visual search of acoustic confusability or interference. Gibson and Yonas provided auditory input while their subjects were scanning down a list of 30 strings of four letters each.[11] The letters presented by earphone sometimes were those selected for low auditory confusability with the target letter (G), whereas in other cases the auditory letters possessed high auditory confusability with the target. High auditory confusability did not slow search, though search was slowed by high visual confusability with the target, and Gibson and Yonas concluded that "these results definitely weaken the hypothesis that visually perceived letters are encoded to acoustic representations as they are scanned" (p. 164).

As Gibson and Yonas conceded, though, greater auditory interference might have been obtained via a visual channel rather than via earphones. Sperling attempted, unsuccessfully, to interfere with his subjects' written reports of the letters in a brief visual display, by turning on auditory noise or a recording of the subject's own voice speaking letters (see n. 1). The auditory interference was very loud (90 db. SPL), yet caused very little decrease in the accuracy of the subjects' reports. Sperling said his subjects could attend selectively to their own rehearsal, prior to reporting the letters, by localizing the rehearsal inside the head and localizing the speech interference at the ears. They also counteracted the auditory noise by more forcefully mouthing and speaking the letters before making the written response.

---

[8] S. Sternberg, Two operations in character recognition: Some evidence from reaction-time measurements, *Percept. & Psychophys.*, 2, 1967, 45–53, 50.

[9] E. R. F. W. Crossman, The measurement of discriminability, *Quart. J. exp. Psychol.*, 7, 1955, 176–195.

[10] U. Neisser and H. K. Beller, Searching through word lists, *Brit. J. Psychol.*, 56, 1965, 349–358.

[11] E. J. Gibson and A. Yonas, A developmental study of the effects of visual and auditory interference on a visual scanning task, *Psychon. Sci.*, 5, 1966, 163–164.

Kaplan, Yonas, and Shurcliff embedded the acoustic confusability in the visual channel, rather than feeding in confusing auditory input via earphones, yet they, too, like Gibson and Yonas, found no effect of acoustic confusability, but a significant effect of visual confusability.[12] However, Kaplan et al. used only two target letters (E and K), and their subjects may have become so adept at looking for key visual features of these two shapes that no further encoding into acoustic representations was necessary.

Chase and Posner used more than two target letters, yet they, too, found no effect of acoustic confusability on visual search and visual comparison.[13] However, in the visual comparison task, the target was present during search and the subjects thus could compare the display shapes directly with the target shape. In the visual search task, the target was present for a considerable duration (10 sec.) prior to search, and a subject thus might have formed strong visual images of the targets, with no need to deal with acoustic representations. Also, Chase and Posner presented the acoustically confusing population of letters (B, C, D, E, P, T, V, Z), from which both display and target letters were drawn, to one set of subjects, and a neutral population to another set of subjects. Even though the two groups of subjects did not differ significantly during the practice day, extraneous factors might have caused those receiving the acoustically confusing set to search faster and thus concealed the effect of acoustic confusability.

Corcoran, on the other hand, has found evidence for acoustic factors in visual search. Corcoran had his subjects cancel the *e*'s in a prose passage, and when data on the *e*'s in the word 'the' were excluded, his subjects proved significantly more likely to miss silent *e*'s than pronounced *e*'s (15% versus 4% errors).[14] Corcoran concluded that "the acoustic image of a word is scanned as well as the visual stimulus" (p. 658). In a second study, Corcoran omitted the *e*'s in certain words and had his subjects indicate where letters were missing. Again, the absent silent *e* was less likely to be detected than the absent pronounced *e*.[15] Corcoran and Weening similarly found that omitted *k*'s that are silent are less likely to be detected than those that are pronounced.[16]

---

[12] G. A. Kaplan, A. Yonas, and A. Shurcliff, Visual and acoustic confusability in a visual search task, *Percept. & Psychophys.*, 1, 1966, 172–174.

[13] W. G. Chase and M. I. Posner, The effect of auditory and visual confusability on visual and memory search tasks, paper presented at the meeting of the Midwestern Psychological Association, Chicago, 1965.

Since Corcoran used words in prose passages, his positive findings do not contradict negative findings from experiments in which subjects searched through nonsense letter sequences. Pronouncing a word is far different from pronouncing the individual letters in an item. Krueger found that his subjects tended to subvocalize more frequently, as indicated by their postsession reports, when searching for a letter through visually presented words rather than nonwords, and through words arranged in coherent, meaningful order rather than in scrambled fashion.[17]

Krueger also found sizable individual differences among subjects on frequency of subvocalization. Some subjects never subvocalized, whereas others almost always 'heard' the visually presented word and 'listened' for the target letter. The effect of acoustic confusability on search for a letter through nonsense items for subvocalizing subjects may have been concealed in the experiments of Gibson and Yonas, Kaplan et al., and Chase and Posner because of a preponderance of nonsubvocalizing subjects.

In the experiments reported below, subjects' reports, obtained after the search session, were used to categorize the subjects as high or low subvocalizers. It was predicted that high subvocalizers would be more slowed by acoustic confusions between target and display letters than low subvocalizers. Since the display letters chosen to be acoustically similar might also be visually similar to the target letter, comparing high and low subvocalizers also allowed the visual factor to be excluded. That is, if visual confusability influenced the results, the same influence presumably should be exhibited by low subvocalizers as by high subvocalizers and thus should not affect the relative effect for high versus low subvocalizers.

---

[14] D. W. J. Corcoran, An acoustic factor in letter cancellation, *Nature*, 210, 1966, 658.

[15] D. W. J. Corcoran, Acoustic factors in proof reading, *Nature*, 214, 1967, 851–852.

[16] D. W. J. Corcoran and D. L. Weening, Acoustic factors in visual search, *Quart. J. exp. Psychol.*, 20, 1968, 83–85.

[17] L. E. Krueger, Search time in a redundant visual display, unpublished Ph.D. thesis, Harvard University, 1969.

Experiments I and II are reported together because they differed only in that the search display, which contained 6 letters in Experiment I, was increased to 18 letters in Experiment II in an attempt to increase the effect of acoustic confusability.

EXPERIMENTS I AND II

Method

Definitions. 'Search time' refers to the duration between onset of the display and *S*'s response. 'Target' refers to the predesignated critical letter for which *S* searches, and 'display' to the row of continuously present visual letters through which *S* searches. A 'same-sounding' display is one whose letters all have a vowel sound which matches that of the target letter (e.g., B D C T V D is a same-sounding display when C is the target). 'Different-sounding' displays contain letters whose vowel pronunciation differs from that of the target (e.g., F S C X L M when C is the target). 'Time savings for different-sounding' indicates how much longer search took, on the average, for same-sounding (acoustically confusing) than for different-sounding displays.

Equipment and manner of stimulus presentation. All letters were capital letters and were illuminated on a display scope by a PDP-4 computer, which also clocked, in milliseconds, *S*'s search time. *S* sat alone, 2–2.5 ft. from the display scope, in a dim room lit by one 25-w. lamp.

Each letter was .19 in. wide and .28 in high. A .28-in. gap separated one letter from the next in the row in the 6-letter displays of Experiment I, while a .19-in. interval separated letters in the 18-letter displays of Experiment II. The target letter appeared .28 in. to the left of a small fixation X; and .75 in. to the right of the fixation point appeared the leftmost letter of the display. A blue-pass filter eliminated a long-persistence yellow phosphorescence of the display scope.

The *S* rested each hand on a separate button, pressing the right one to indicate the target was present and the left to indicate it was absent. When a button was pressed, an electric circuit was closed and the clock stopped, and the computer read and recorded the elapsed search time.

Procedure. The *S*s were instructed to search as fast as possible, and to press the 'yes' button as soon as they saw the target letter in the display, but to avoid making any errors. They were told to look at the X fixation point while the target was on and not to shift their gaze to the display until it actually appeared.

The target letter appeared for .4 sec., after which ensued a 1.0-sec. 'blank' period, during which only the fixation point was visible, before the display appeared. *S*'s response extinguished the display, and the screen was blank for .8 sec. before the target for the next trial appeared. The target never appeared more than once in the display and did not appear at all on one-fourth of the trials. When *S* pressed the incorrect button, the word 'WRONG' appeared on the screen for 2.0 sec.

After the search session, *S* was questioned on how he went about the search task in general. He also was asked how often he pronounced, either aloud or to himself, the target letter when it appeared, and how often he pronounced letters in the display.

Display materials. Two types of displays were presented: displays with letters whose pronunciation ends in ē, and those with letters whose pronunciation begins with ĕ. To create an ē display of 6 letters (Experiment I) or 18 letters (Experiment II), the computer sampled from a set of 8 letters (B, C, D, G, P, T, V, Z). Sampling was random and with replacement, but with the constraint that the same letter not be shown twice in a row. Sampling also was constrained so that the target letter was not included. To create an ĕ display, sampling was from a pool of 6 letters (F, L, M, N, S, X). If the target letter was to be present, it replaced 1 letter in the display so created.

A possible visual artifact is that those displays which provided an auditorily more confusing context for the embedded target letter also may have provided a visually more confusing context. That is, the display letters that sounded more like the target letter may also have looked more like the target. However, confusion scores from a legibility test by Hodge[18] indicate that when one of the 14 capital letters (8 ē and 6 ĕ) used in the present experiments is presented, *S* is more likely to report (incorrectly) one of the six or eight different-sounding letters (.11% of all trials) than one of the six or eight same-sounding letters (.06%). Thus, the same-sounding displays presented an acoustically more confusing context, but seemingly not a visually more confusing context.

Experimental design. Each *S* received all conditions and thus served as his own control. Eight replications of a basic set of 32 conditions yielded 256 trials in all for each *S*. (Within a set of 32 trials, the various conditions were sampled at random without replacement by the computer.) The 32 conditions represent all combinations of four types of target-display combinations and eight target locations. The eight locations consisted of two trials in which the target was not present; and six trials in which the target appeared once in each of six letter positions, which were 1, 2, 3, 4, 5, and 6 in the 6-letter display of Experiment I, and 3, 6, 9, 12, 15, and 18 in the 18-letter display of Experiment II.

The four target-display combinations consisted of two same-sounding (ē target and ē display, ĕ target and ĕ display) and two different-sounding combinations (ē target and ĕ display, ĕ target and ē display). Thus, on half

[18] D. C. Hodge, Legibility of a uniform strokewidth, *J. engr. Psychol.*, 1, 1962, 34–46.

the trials the display contained letters having the same vowel sound as the target letter.

Data analysis. On search time, a geometric mean was taken for the eight replications of a particular condition. When *S* pressed the wrong button on one of the eight trials, that trial was removed and averaging performed on the remaining trials. All further averaging was arithmetic, and for each *S* a (mean) same-sounding search time was obtained by averaging the geometric means for the 16 same-sounding conditions, and a different-sounding search time by averaging the 16 different-sounding conditions. Since each *S* received all conditions, a single figure of merit (i.e., time savings for different-sounding displays) was obtained for each *S* by subtracting different-sounding search time from same-sounding search time, and provided the data for a single-mean t test. A positive figure on time savings for different-sounding displays indicates, as predicted, that search time was longer when the display letters had the same vowel sound as the target, and thus were acoustically confusing. Lastly, since search-time data were excluded for trials in which *S* pressed the wrong button, errors provide an independent measure of the difficulty of search through same-sounding and different-sounding displays. Because of their greater variability, errors were analyzed by the sign test.

Subjects. The *S*s were 10 Harvard and Radcliffe undergraduate and graduate students in Experiment I, and 20 in Experiment II.

Results

Table I shows that in both experiments, search time was longer, as predicted, for the same-sounding condition (in which display letters had the same vowel sound as the target letter) and that as the target was placed farther toward the right in the display, the difference between same-sounding and different-sounding displays generally increased. The 23-msec. difference in search time between same-sounding and different-sounding displays (823 versus 800 msec.) in Experiment I was only marginally significant [t (9) $= 1.62$, $p < .10$, one-sided test], as was the 25-msec. difference (1,758 versus 1,733 msec.) in Experiment II [t (19) $= 1.69$, $p < .10$, one-sided test]. When the data from the two experiments are combined, the longer search time for the same-sounding condition was significant [t (29) $= 2.25$, $p < .025$, one-sided test].

Similarly, the slightly higher error rate for same-sounding than for different-sounding displays (5.3% versus 4.5%) in Experiment I was only marginally significant (sign test, $p = .09$, for 7 positives, 2 negatives, 1 tie), as was the difference (5.4% versus 5.2%) in Experiment II (sign test, $p = .11$, for 11 positives, 5 negatives, 4

TABLE I

Search Time by Letter Position of Target, for Same-Sounding and Different-Sounding Displays

Letter position of target	Same-sounding (msec.)	Different-sounding (msec.)	Time savings for different-sounding (msec.)
		Experiment I	
1	724	737	−13
2	734	710	+24
3	662	671	−9
4	737	680	+57
5	792	762	+30
6	855	798	+57
Target not present	1,039	1,020	+19
All trials	823	800	+23
		Experiment II	
3	958	944	+14
6	1,136	1,121	+15
9	1,406	1,402	+4
12	1,639	1,640	−1
15	1,930	1,807	+123
18	1,908	1,863	+45
Target not present	2,545	2,546	−1
All trials	1,758	1,733	+25

Note: In averaging, trials in which the target was not present received twice the weight of a regular letter position, reflecting their greater frequency during the session. $N = 10$ for Experiment I, $N = 20$ for Experiment II.

ties). Again, when the data for the two experiments are combined, the difference was significant (sign test, $p = .02$, for 18 positives, 7 negatives, 5 ties).

Only 5 of the 30 subjects in the two experiments reported ever pronouncing display letters, but all but 3 subjects reported pronouncing at least some of the target letters, and 10 subjects did so on virtually all (90% to 100%) of the trials. On the average, the subjects in Experiment I pronounced 60% of the targets, and the subjects in Experiment II 58%.

Further, there was a marked association between frequency of reported pronunciation of target letters and time savings for different-sounding displays. In Experiment I, pronunciation of target letters correlated .11 with time savings for different-sounding displays. When the very discrepant results on 1 subject were removed, the correlation was .58. In Experiment II, the correlation was .36, and for the 30 subjects in the two experiments combined the correlation was .30.

Since the correlations were not significant, a further test was made by dividing the subjects into two equal groups: high pronouncers and low pronouncers of target letters. In Experiment I, the 5 high pronouncers, who reported pronouncing 80% to 100% of the target letters, showed a 30-msec. time savings for different-sounding displays, compared with 16 msec. for the 5 low pronouncers. The difference between high and low pronouncers was not significant [t (8) = 1.11, $p = .15$, one-sided test]. In Experiment II, the 10 high pronouncers (75% to 100%) showed a 53-msec. time savings for different-sounding displays, compared with −4 msec. for the 10 low pronouncers, and the difference was significant [t (18) = 5.55, $p < .0025$, one-sided test]. The difference between the 15 high and the 15 low pronouncers in the two experiments combined also was significant [t (28) = 4.88, $p < .0025$, one-sided test]. On errors, high pronouncers did not differ from low pronouncers: 9 of the 15 high and 9 of the 15 low pronouncers had more errors on same-sounding than different-sounding displays.

The subjects who more frequently pronounced the target letters were probably more prone to encode more fully and form acoustic representations of display letters, even though they denied pronouncing display letters; they would thus be expected to search slower in general. Search time on different-sounding displays, which is taken to indicate general search speed, correlated .58 with frequency of pronunciation of target letters in Experiment I and .03 in Experiment II. Average search time was 1,484 msec. for the 15 high pronouncers, compared with 1,385 msec. for the 15 low pronouncers.

Discussion

Visual search was faster and more accurate through letters that did not sound similar to the target letter than through letters that did, which indicates that subjects do form acoustic representations of display letters in visual search. That the results reflect acoustic factors, and not some visual artifact, is indicated by the greater effect for high than for low pronouncers of target letters.

The (mean) 45-msec. time savings for different-sounding displays among the 15 high pronouncers understates the potential acoustic effect, because it includes trials in which the subjects did not pronounce the target letter. If electromyograms had been taken, so that just those trials on which subjects actually subvocalized or pronounced the target could be sifted out, acoustic confusability might show a much larger effect. Electromyograms also would clear up the question of whether the subjects actually subvocalized some display letters, which is implied by the slower search through same-sounding than different-sounding displays, but which the subjects denied doing. In any case, the subjects did not search so fast that they could not have been subvocalizing display letters, because Landauer reports that it takes only 3.5 sec. to say aloud the English alphabet,[19] whereas the subjects in Experiment II took about 2 sec. to search through 18 letters (see Table I).

That the subjects' reports of target-letter pronunciation, but not display-letter pronunciation, were associated with the effect of acoustic confusability indicates that a necessary, but hitherto overlooked, condition for demonstrating acoustic confusions in visual search is that the target letter itself be encoded (by the subjects) primarily as an acoustic representation. That is, even if display letters are fully encoded as acoustic as well as visual representations, acoustic confusability may have no effect if the target letter is encoded primarily as a visual representation. Keeping one target letter in mind during visual search puts a minimal load on short-term memory, and subjects may normally rely on their memory of the shape rather than the sound of the target letter. If the subjects were to search for several letters at a time, however, remembering the targets would become much more difficult, as Krueger found, and the subjects might revert to auditory storage, which Murray has shown to be more efficient than visual storage; and acoustic confusability would have a larger effect.[20]

Unexpectedly, lengthening the display from 6 letters (Experiment I) to 18 letters (Experiment II) did not increase the effect of acoustic confusability. Placing the target farther to the right generally increased the time savings for different-sounding displays, and increasing the display length ought to have had the same effect, yet the time-savings figures are virtually identical for Experiments I and II (23 versus 25 msec.). One possible explanation is that since only 8 *ē* and *ě* letters were drawn upon, most letters were repeated several times in an 18-letter display, giving the subjective impression (which this investigator experienced) that the display recycles every 5 or 6 letters, and making the target letter, when it does appear, stand out visually as the only letter which has not been seen repeatedly. Further, when letters are presented repeatedly, progressively less processing may suffice to discard them at each reappearance, and the subject may not continue to encode auditory features of a recurring letter; so again less effect of auditory confusability would be expected with a longer visual display. The negative findings on acoustic confusability of Kaplan et al. (n. 12) similarly may be due to the repetition of letters, because they used only 4 different letters in constructing a given display of 30 4-letter items. For example, one of their displays contained only the letters C, G, I, T.

---

[19] T. K. Landauer, Rate of implicit speech, *Percept. mot. Skills,* 15, 1962, 646.

[20] Krueger (n. 17); D. J. Murray, Vocalization-at-presentation, auditory presentation, and immediate recall, *Nature,* 207, 1965, 1011–1012.

CORRECTION PROCEDURES IN THREE-ALTERNATIVE VERBAL-DISCRIMINATION LISTS WITH INTERTRIAD ASSOCIATIONS

JOHN H. MUELLER and JAMES P. PICKERING,
University of Missouri, Columbia

Abstract. Intertriad associative relationships were varied in three-alternative verbal-discrimination lists, using either specific correction (the word itself) or outcome correction ('right,' 'wrong'). All of the associative lists were learned faster with outcome-correction feedback than was the control list, in which there were no interpair associates (NA list); but with specific-correction feedback, only the list in which both associates were right (AR list) was learned faster. The facilitation with outcome correction is attributed to a change in the chance level of guessing for the lists in which both associates were wrong (AW list) or in which the associates were both right and wrong in different pairs (AB list), and to a summation of rehearsal and implicit response in the AR list. The failure to find interference with specific correction for the AW and AB lists seemed to be a result of the ceiling effect that occurred because of the slow presentation rates.

Verbal-discrimination learning has recently been explained in term of the subject's ability to distinguish between the past frequency of experience of the right (R) and wrong (W) members of each pair.[1] It was assumed that these differences in experience were acquired through four types of events which contributed one 'frequency unit' each time they occurred during the course of learning. These events were the act of perceiving an item, or the *representational* response; the choice or *pronunciation* response during anticipation; the *rehearsal* of the correct response during the feedback interval; and the *implicit* associative response. The rehearsal response was seen as the event initially shifting the frequency of experience in favor of the R term, because the representational and

Received for publication May 18, 1970. The investigation was supported by Biomedical Sciences Support Grant Fr-07053 from the General Research Support Branch, Division of Research Resources, Bureau of Health Professions Education and Manpower Training, National Institutes of Health.

[1] B. R. Ekstrand, W. P. Wallace, and B. J. Underwood, A frequency theory of verbal-discrimination learning, *Psychol. Rev.*, 73, 1966, 566–578.

pronunciation responses could contribute only at random if the R and W terms were of comparable strength to begin with. Thus frequency theory assumed that the subject's rehearsal of the correct response during feedback eventually led to a difference in the number of times that each R term had been experienced relative to its W term, so that the subject could perform correctly by simply following a 'rule' and always choosing the most familiar member in each pair (Rule 1).

The implicit-response unit was considered to provide an increment in experience whenever the implicit response to one item occurred explicitly as a member of another pair (e.g., *king-x* and *queen-z* as pairs). It has been found that performance was facilitated if the associates both occurred as R terms (AR), but was hindered if the associates were both W terms (AW) or if the associates served both (AB) as a W term and an R term in different pairs.[2] This was expected by frequency theory, because the implicit response would combine with the rehearsal response in the AR list, would offset it in the AW list, and would largely obviate the use of a consistent rule in the AB list (with the 'most frequent' appropriate for half of the pairs and the 'least frequent' for the other pairs). When the associates were members of a second list, these effects were most pronounced for those subjects instructed about the relationships.[3] With intrapair associates (e.g., *king-queen* as a pair), the evidence is mixed, as some studies find this pairing is harder than a control pairing,[4] while other investigations find no difference relative to unrelated pairings.[5]

The present experiment attempted to further test the operation of the implicit-response unit within the frequency-theory framework, with the effect of the rehearsal response considerably reduced. Theoretically, the implicit response in an interpair associative arrangement (e.g., an AR list) could initiate the frequency shift toward the R terms in the manner that the rehearsal response is assumed to do in lists with no interpair associates (NA), as described above. A rigorous test of this proposition is probably impossible if the implicit response can function in the manner proposed, since differences induced by these implicit responses will lead to above-chance performance. The problem is to somehow inform the subject of the outcome of his choice without allowing the rehearsal response to occur prior to the discovery of the correct response.

---

[2] Ekstrand, Wallace, and Underwood (n. 1 above).

[3] D. C. Raskin, C. Boice, E. W. Rubel, and D. Clark, Transfer tests of the frequency theory of verbal discrimination learning, *J. exp. Psychol.*, 76, 1968, 521–529.

[4] D. S. Palermo and J. R. Ullrich, Verbal discrimination learning as a function of associative strength between the word-pair members, *J. verb. Learn. verb. Behav.*, 7, 1968, 945–952.

[5] F. E. Fulkerson and D. H. Kausler, Effects of intrapair and interpair bidirectional associates on verbal-discrimination learning, *J. verb. Learn. verb. Behav.*, 8, 1969, 307–310.

Feedback in verbal-discrimination learning has been investigated previously,[6] but not with procedures intended to reduce the rehearsal response. In the present experiment, some subjects were told only 'right' or 'wrong' after each response; this was the outcome-correction condition. Other subjects were in the specific-correction condition; here R terms were designated during the feedback interval, as is commonly done by underlining. Two-alternative lists would always provide enough information for the rehearsal response to occur with either procedure, so three-alternative lists of the AR, AW, AB, and NA constructions were used. Even with this procedure the subject could still randomly choose from the two remaining alternatives for rehearsal when wrong in the outcome-correction condition, so further control was sought. The subjects with outcome-correction feedback were required to pronounce each of the remaining alternatives when wrong, thus distributing at least one frequency unit to each of the three alternatives on an error trial via pronunciation responses. A procedure somewhat like this has been used with two-alternative NA lists.[7] It should be noted that this will not completely eliminate the rehearsal response, although it will not be operative for each triad from the very first trial as is the case with specific-correction feedback. With the outcome-correction feedback, the rehearsal responses can operate following correct guesses, and guessing would be affected by the differences induced by the implicit responses, as explained below.

Certain predictions follow from frequency-theory considerations for lists with interpair associates under the different feedback conditions, but these are most readily explained in the context of specific examples, as presented in Table I. The two basic lists used in this study are shown, and the four list conditions were derived from them as follows. The three left-hand columns contain all of the triads for the AR, AW, and AB constructions. The associates are listed in the third column; these were the R terms for AR, with the other two words serving as W terms. The designated items in the second column served as R terms in the AB list, with the corresponding associate becoming one of the W terms, and retaining the other associates as R terms. In the AW arrangement, all of the second-column words served as R terms, with the associates always W terms. For the control, or NA, construction, the words in the right column simply replaced the associate immediately to their left to eliminate the interpair associations, with the third column the R terms. The basic materials were adapted from Ekstrand, Wallace, and Underwood (n. 1). It should be noted that all four groups have a common set of W terms, the words in the first column.

---

[6] J. T. Spence, The effects of verbal reinforcement combination on the performance of a four-alternative discrimination task, *J. verb. Learn. verb. Behav.*, 6, 1966, 421–428.

[7] D. H. Kausler and R. J. Sardello, Item recall in verbal-discrimination learning as related to pronunciation and degree of practice, *Psychon. Sci.*, 7, 1967, 285–286.

It would be predicted that a control list without implicit responses (NA; formed by replacing the appropriate words in column three) would be rather difficult to learn with outcome-correction feedback, since chance (⅓) discovery would be the only way in

TABLE I

WORD LISTS USED

Associative triads			Control words
bird	head	chair	
foot	rain[†]	table	carpet
rough	clear[†]	house	
print	bread	home	dark
fast	cold	queen	
short	pepper[†]	king	door
acorn	sweet[†]	needle	
shallow	flower	thread	ugly
hammer	web	road	
doctor	curl[†]	street	round
lion	smoke	hill	
baby	metal[†]	mountain	white
bottom	soft[†]	steal	
apple	flag	thief	horn
wind	low[†]	town	
cat	brand	city	bed

[†] Correct in the AB list.

which the R terms could be found. Once found, however, the rehearsal response could occur. It might also be expected that such lists would involve a substantial amount of associative learning as well, with the subject using the association between the R term and the W terms as well as frequency cues.[8]

In an AR list, the presence of implicit responses should speed discovery, since outcome-correction feedback equalizes experience for triad members when the subject is wrong, with the implicit response occurring to benefit another triad. For example, even if the subject chooses *head* in the first triad in Table I, he must pronounce *chair* as well, so that the implicit response *table* can occur. When the second triad occurs later, *table* has by then been experienced more than *foot* or *rain*, and is more likely to be chosen than would be expected by chance alone. Assuming that the implicit- and rehearsal-response units contribute equally in terms of experience, a group given outcome correction on AR lists might compare favorably to a group given specific correction on NA lists, and could perform better than a group given specific correction on NA lists because the rehearsal response is not completely eliminated with their outcome-correction feedback.

For the AW construction, the implicit response contributes to the W term. This has been shown to produce interference relative to a NA list with two alternatives (see n. 1). Frequency theory explains this by noting that the implicit response defeats the rehearsal response. When the subject chooses *chair, table* receives an increment, so that even if he correctly chooses *rain* later and rehearses it, that rehearsal only provides one unit to offset the implicit response to the W term in that pair. The result is that *rain* and *table* remain equal in two-alternative lists. With three-alternative lists, however, the outcome-correction procedure actually allows the implicit response to provide an extra increment, since the pronunciation responses on error trials equalize experience otherwise. That is, the rehearsal response does not always occur to offset the implicit response to the W term, so that the associate W term may maintain the extra frequency at the end of the trial. If the subject then chooses by avoiding the most frequent, he increases his chance of discovery from ⅓ to ½. For example, assume that the third column in Table I is wrong and the second column correct (AW). If the subject chooses *chair,* this also adds via an implicit response to *table.* Subsequent choices may involve avoiding the more frequently experienced items (Rule 2), so that the actual decision is between *bird* and *head* or *foot* and *rain* rather than the nominal triads. Eventually, the subject will discover the R terms, and rehearsal responses will increase the experience for the R terms to match that of the associate W terms. At this point, some confusion may result as the subject is now faced with two equally likely items (e.g., *head* and *chair*) but must switch to Rule 1. This may be compensated for by the fact that the probability of guessing is again above chance (⅓).

---

[8] D. H. Kausler and J. A. Boka, Effects of double functioning on verbal discrimination learning, *J. exp. Psychol.*, 76, 1968, 558–567.

Therefore, the AW triads may only involve *pairs* at the functional level, if frequency theory is correct. The same argument applies to half of the AB triads also. As a result, the interference obtained with two-alternative lists may not obtain for AW and AB triads with outcome-correction feedback. Furthermore, if this interpretation is correct, frequency theory might predict that the AW and AB constructions would produce more errors involving the use of the associate W terms, rather than the common unrelated W term (column one in Table I). However, this prediction is somewhat dependent upon the existence and the duration of the two stages just described—upon how long the subject spends choosing between *bird* and *head* (Rule 2) rather than between *head* and *chair* (Rule 1). If these phases are of equal duration, which seems unlikely, there would be no difference expected. If the latter stage is longer, errors should increasingly involve the associate W term. If so, then the recall of the two types of W terms should also favor the associate W term.

METHOD

Design and subjects. A 2 × 4 factorial was used, with type of correction procedure—specific correction (SC) and outcome correction (OC)—and type of list (AR, AW, AB, NA) as the two factors. Five men and five women, students in introductory psychology courses, served as the *Ss* in each of the eight groups in partial fulfillment of course requirements.

Materials. As described previously (Table I), two basic lists of 16 triads were constructed. One served as the NA list by substitution of the control words in column four for one of the associate words in column three. The first three columns as shown in Table I served as the triads for all of the associate groups, with *E*'s response during feedback determining AR, AW, or AB. Feedback was verbal, with *E* stating the specific word for the specific-correction condition, and 'right' or 'wrong' for the outcome-correction condition, so that the same tapes served for both conditions. Both lists were arranged in four different orders for presentation, with each member of a triad appearing in the left-middle-right spatial position at least once over these orders, and with the disparity in spatial position of the R term balanced over the orders. Each order was used as the first order approximately equally often.

Procedure. The lists were presented on a Lafayette drum at a 4:4-sec. rate, with an 8-sec. intertrial interval. These longer intervals were used to accommodate the post-feedback performance of the *S*s given outcome-correction feedback.

In the specific-correction condition, the triad appeared for 4-sec., during which time the *S* responded, with the triad reappearing in the drum window and *E* verbalizing the correct member each time. In the outcome-correction condition, the anticipation was the same, but when the triad reappeared the *S* was merely told the outcome of his choice by *E*, either 'right' or 'wrong.' If incorrect, the *S* was instructed to repeat aloud each of the remaining words once, thus matching the pronunciation response to the other alternative during the anticipation frame. If correct, the *S* given outcome correction was instructed to repeat the term he had just chosen, matching the statement of the R term by *E* during feedback for the specific-correction condition. The *S*s with outcome-correction feedback should experience the same distribution of frequency as the *S*s with specific-correction feedback on correct anticipations, but the outcome-correction procedure on error trials should largely limit the effects to implicit responses. The *S*s were not informed of the intertriad associative relationships.

Performance was continued to two perfect trials, and was followed by an unpaced W-term recall test. The *S*s were given a sheet of paper with the R terms in one column and two blanks beside each for the appropriate W terms. A space was provided at the bottom for additional W terms which the *S* could not match to any R term, and these were considered in a lenient scoring.

RESULTS

The number of trials to successive criteria (8/16 to 16/16) and the number of trials to list criterion (two consecutive perfect trials) are shown in Figure 1. For a 2 × 4 factorial analysis of variance, the trials to criterion indicated that the specific-correction feedback was superior to the outcome-correction feedback [F (1, 72) = 182.89, $p < .001$], and this superiority was significant from the 6/16 criterion to the list criterion in separate analyses. Trials to list criterion also indicated an effect for list type [F (3, 72) = 11.90, $p < .001$], which was significant as early as 14/16 in separate analyses. The interaction of correction procedure with list type was marginally significant in the list-criterion data [F (3, 72) = 2.42, $p < .10$], with the greatest effect of list type in the outcome-

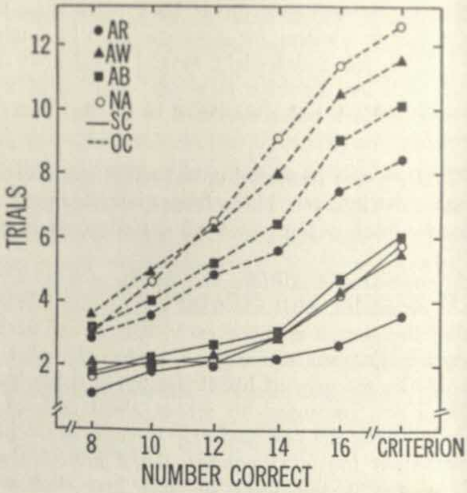


FIG. 1. Number of trials to successive criteria and list acquisition by list type and correction procedure.

correction conditions. Although any effect of the associates seemed largely obscured by the ceiling effect with specific correction and the slow rates, the subjects in this condition were significantly superior with the AR list over the NA list ($p < .05$) by individual comparisons (Duncan's multiple-range test) in the list-criterion data.

Most of the expected effects of outcome correction were present. The subjects in this condition were not handicapped on list AW relative to NA, as they presumably benefited from the change in guessing probability mentioned earlier, but the apparent facilitation of AW relative to NA did not reach significance ($p > .10$). In the AB list, the gain combined with the early use of Rule 1 for the AR half of the list to yield a significant gain over both NA and AW ($ps < .01$ and $.001$ respectively) in the list-criterion data. Facilitation for AB and AW is contrary to previous findings with specific-correction feedback and two-alternative lists (see n. 1), but seems to follow from the frequency-theory considerations noted previously, assuming a change in guessing probability.

Comparability of the group given outcome correction on the AR list and the group given specific correction on the NA list was not found, with the difference shown in Figure 1 significant from 10/16 to list criterion ($p < .01$). The group given outcome correction on the AR list also differed from all other outcome-correction groups as well. The comparison between it and the group given specific correction on the NA list may indicate that the rehearsal response and the implicit response are not equivalent in terms of the magnitude of subjective experience provided, or it may indicate that the implicit response is less likely to be repeatedly evoked during a single feedback frame. This is a difficult comparison here anyhow, since the outcome-correction groups do get rehearsal responses also following R-term discovery. It might also be expected that the difference between the outcome-correction groups on the AR and NA lists would be about the same as the difference between the specific-correction groups on the AR and NA lists, but that was not the case either. The effects of implicit (associative) responses were present with outcome-correction feedback, but the difference in the pronunciation responses for the two feedback conditions as administered here confounds any statement about the equality of implicit responses and rehearsal responses.

Table II presents the percentage of all errors that involved the use of the associate W term for the AB and AW groups with both types of feedback. For the AW group, 50% would be chance, while 25% would be chance in the AB groups. It had been expected that the associate would occur as an error more often than the unrelated W terms, but only the group given specific correction on the AB list showed a reliable difference from chance (37.2). Considering the representation of the associate W terms in total W-term recall (leniently scored), it seems that both AB groups were somewhat above the chance level. The prediction of differential usage of the associate W term received rather weak support, especially in that the W-term recall did not always covary with the use of the associate as an error; that is, the group given specific correction

TABLE II

PERCENTAGE OF ERRORS INVOLVING THE ASSOCIATE W-TERM AND PERCENTAGE OF TOTAL W-TERM RECALL THAT WAS THE ASSOCIATE W-TERM

	Specific correction		Outcome correction	
Measure	AW	AB	AW	AB
Errors	49.7	37.2	48.4	27.3
Recall	53.5	29.1	50.1	31.3

on the AW list indicated somewhat greater recall of the associate, but not greater use of the associate in errors.

Pooling all W-term recall without regard to type or correct pairing (Table III) indicated a significant effect for the correction procedure [F (1, 72) = 13.07, $p < .001$]. List type also produced a significant effect [F (3, 72) = 5.35, $p < .01$] and interacted significantly with correction procedure [F (3, 72) = 5.22, $p < .01$] due to the performance on AB and NA. While some of the recall effects may merely reflect the degree of list learning, the superior performance of the group give outcome correction on the AW list does not seem readily attributable to such an artifact.

TABLE III

Total W-Term Recall, Pooled Over Type and Scored without Regard to Correct Pairing

Feedback	List construction			
	AR	AW	AB	NA
Specific correction	5.8	8.6	11.7	11.1
Outcome correction	10.0	19.7	11.3	12.3

DISCUSSION

The acquisition data seemed in accord with the predictions, although there was not convincing evidence for the differential use of associate W terms as errors. While the ceiling effects with specific-correction feedback do not allow the complete extension of all findings from two-alternative lists (see n. 1), the implicit responses did seem to facilitate the AR list in the specific-correction condition. It seems reasonable to attribute the absence of interference for AW and AB lists in the specific-correction conditions to the slow presentation rates used to accommodate the outcome-correction groups. With outcome-correction, however, the failure to find interference with AW and AB lists is consistent with the notion of a changed level of guessing, with the subjects functionally choosing between the R term and the associate W term rather than all three alternatives. This change in guessing level seemed to more than compensate for any interference, especially in the AB arrangement.

The predictions based on the implicit-response (IAR) mechanism were not as well supported in the error and recall data. Any evidence supporting those predictions would also be consistent with the assumption that more associative learning was occurring in those groups (see n. 7). It could be assumed that an incidental association is formed in verbal-discrimination learning in addition to differential frequency accumulation. Such associations would be expected to be utilized more when rules could not be used (e.g., AB) or with outcome-correction feedback. This would also lead to an increase in W-term recall, as was, in fact, present for the groups given outcome correction (see Table III), and might not entail differential usage of the associate as an error. It must be noted that the error-usage prediction was not certain but rather a function of the duration of the two performance stages assumed to be occurring. The present data do not provide a clear choice between the two explanations. However, the multiple-alternative verbal-discrimination list would seem to be useful for the further study of the operation of the implicit-response unit in intertriad and intratriad associative lists.

RECALL OF A WORD LIST AFTER INTERPOLATED READING

ISABEL M. BIRNBAUM, *University of California, Irvine*

Abstract. The study investigated the effects of natural language activity on the retention of a word list. After learning a paired-associate list of high-frequency words, subjects in a control group engaged in a pattern-selection task while subjects in three experimental groups read descriptive materials. Half of the words from the list were embedded in the reading material. Accuracy and latency of recall of the word list was measured 21 min. after the end of original learning. There was no difference between groups in the number of words correctly recalled; in all three experimental groups, however, there was a significant increase in the latency of recall of responses only for pairs which had been embedded in the interpolated reading material. The conclusion is that natural language activity can have specific interfering effects on the recall of a word list.

The fact that naive subjects show retention losses for a single, laboratory-learned list continues to pose explanatory problems for an interference theory of forgetting.[1] The assumption that proactive inhibition from prior language habits accounts for retention loss has received little support. It has been suggested that the forgetting of a single list over uncontrolled retention intervals may reflect retroactive inhibition from language activity rather than proactive inhibition.[2] In brief, it is assumed that the A–B, C–D transfer paradigm is formed by the laboratory list (A–B) and subsequent language activity (C–D) during the retention interval. (There are different, unrelated words in the two situations.) A reduction in the availability of first-list responses (contextual un-

---

Received for publication March 17, 1970. This research was supported by Research Grant MH-13629 from the National Institute of Mental Health, United States Public Health Service. The author wishes to thank Joellen Eichner for her valuable assistance in all phases of the research.

[1] L. Postman, Does interference theory predict too much forgetting? *J. verb. Learn. verb. Behav.*, 2, 1963, 40–48; B. J. Underwood and B. R. Ekstrand, An analysis of some shortcomings in the interference theory of forgetting, *Psychol. Rev.*, 73, 1966, 540–549.

[2] G. Keppel, Retroactive and proactive inhibition, in T. R. Dixon and D. L. Horton (eds.), *Verbal Behavior and General Behavior Theory*, 1968.

learning) has been found when lists conform to the A–B, C–D paradigm.[3]

A second possibility is that specific associative loss (a reduction in the tendency of item A_1 to elicit item B_1, item A_2 to elicit item B_2, etc.) may occur when words from the list appear in one or more contexts outside of the laboratory. For example, suppose that one of the pairs (A_1–B_1) in the original list is *minute-story*. These words might occur in sentences written, spoken, or heard during the retention interval. Each word, when it occurs, is likely to appear with words other than its partner from the list, thus forming the A–B, A–C, A–D, and so on, paradigm. For example, the word *minute* might occur during the retention interval in phrases such as 'in a minute,' 'a minute later,' and 'Do you have a minute to spare?' The occurrence of list words during the retention interval may serve to weaken the associative connection between specific words on the list.

The purpose of the present experiment was to explore some implications of a retroactive-inhibition theory of extraexperimental interference. A paired-associate list was learned by all subjects and was followed by an interpolated 'cover task' for the experimental groups. The task required that the subject read a large number of short passages (interpolated reading). The passages were descriptions of college courses and travel tours, written specifically for the experiment. Words from the paired-associate list were embedded in the passages. Two experimental groups were informed of the occurrence of list words in the passages, and one experimental group was not. A within-subjects design was used such that the words from only half of the pairs on the paired-associate list appeared in the passages read by each subject. A control group learned the paired-associate list and then engaged in a pattern-selection task during the retention interval. Recall of the list was measured on an unpaced test, and response latencies were recorded. An overall difference in recall in the experimental versus control groups would indicate the occurrence of contextual interference, a general loss in availability of all responses on the list. In the experimental groups, however, a greater decrement in the recall of pairs whose members were present during interpolated reading would indicate specific interference, a decrement which is contingent upon the occurrence, in the interpolated passages, of words from the list of paired associates. The results would be useful in evaluating the assumption that language activity during uncontrolled retention intervals leads to the forgetting of laboratory-learned lists of words. To this writer's knowledge, the effect of natural language activity on the retention of verbal lists has not previously been explored.

---

[3] J. B. McGovern, Extinction of associations in four transfer paradigms, *Psychol. Monogr.*, 78, 1964, No. 16 (Whole No. 593).

METHOD

Materials. The same 10-pair paired-associates list was learned by all *Ss*. The stimulus and response terms were high-frequency nouns selected from the Thorndike-Lorge word count.[4] Each of the words was rated 'M' in the Lorge magazine count, indicating that it had occurred 1,000 or more times per 4.5 million words. In the set of 20 words, 12 initial letters occurred once, and 4 initial letters occurred twice. The list of paired associates was constructed by a random pairing of the words. There were four different orders of presenting the 10 pairs. The pairs were divided into two sets of 5, designated Set A and Set B. Since estimates of the relative difficulty of the pairs were available from a previous study (unpublished), the sets were equated as nearly as possible in overall difficulty. The pairs in Set A were *trouble-hundred, garden-question, minute-story, morning-dinner, building-color*. The pairs in Set B were *baby-reason, water-order, party-letter, nothing-figure, table-people.*

A separate booklet of interpolated passages was prepared for the 10 words in Set A and for the 10 words in Set B. The passages in each booklet consisted of 20 course descriptions and 20 tour descriptions. The former were like the course descriptions in college catalogs, and the latter were descriptions similar to those which appear in travel brochures. The following examples illustrate the types of passages which were used:

"*Music: Modern Trends.* Some people have said that there has been nothing new in music for over a century. Modern music will be analyzed from the standpoint of historical versus innovative influences."

"*Castle on the Rhine.* The group will spend three days and two nights at this castle of Roman origin. The grounds include the old castle and one new building for guest accommodations. Dinner is served by candlelight each evening in the grand salon of the castle, built in 1280. Visit the ancient vineyards, the formal garden, and historic churches nearby."

In writing the description, 2 or 3 words (stimuli or responses) from the set of five pairs were embedded in each description. For example, in the first passage above the words from the paired-associates list were *people* and *nothing*. Each of the 10 words in the set was used a total of 10 times throughout the passages. A typical description contained two, three, or four sentences, and in no case were both members of the same pair used in the same description. The descriptions were arranged in an order such that all 10 words from the set appeared once before any of the words appeared again, and so forth. This procedure guaranteed that each *S* would come across each word approximately the same number of times, regardless of how far he had proceeded in the booklet when interpolated reading was terminated.

---

[4] E. L. Thorndike and I. Lorge, *The Teacher's Word Book of 30,000 Words*, 1944.

Design. Four groups learned the list of paired associates to a criterion of 7/10 correct responses on one trial. Three experimental (E) groups were then given the task of reading and rating the descriptions. The control group (C) engaged in a pattern-selection task: *S* received a booklet in which 135 sets of five figures were drawn. His task was to indicate which of the figures in each set of five was different from the others. The figures in each set differed in such things as spatial distribution of objects, or number of objects, and so on. The interpolated activity was terminated for all groups after 20 min. and was followed by an unpaced test of recall of the list of paired associates.

The three experimental groups differed on two bases: the instructions for the interpolated task, and the rate of presentation of the paired-associate list. The *S*s in Group E-R (Read) were instructed to read the passages and rate each one on a seven-point scale of personal interest. No mention was made of the occurrence of paired-associate words in the passages. Groups E-RS$_1$ and E-RS$_2$ (Read and Search groups) received the same instructions as did Group E-R and were told, in addition, that some of the words from the paired-associate list would appear in the passages. The *S* was to look for and underline those words from the list which he came across while reading and rating the passages. The distinction between E-RS$_1$ and E-RS$_2$ was that the latter received the paired-associate list at a slower rate of presentation than did all other groups: a 4:5-sec. rate rather than a 2:2-sec. rate was used in Group E-RS$_2$. This group was run in order to investigate the possibility that greater opportunities for natural language mediation during original learning would lead to greater interference from interpolated reading and thinking about the words from the list. Of course, any differences in recall between E-RS$_1$ and E-RS$_2$ would only be suggestive, since the groups differ on a number of other factors such as degree of original learning and similarity of original learning to the unpaced test of recall.

Procedure. The list of paired associates was presented on a Stowe memory drum at a 2:2-sec. rate with a 4-sec. intertrial interval for Groups C, E-R, and E-RS$_1$; a 4:5-sec. rate with a 9-sec. intertrial interval was used for Group E-RS$_2$. Anticipation trials continued until *S* had reached a level of 7/10 correct responses. The interpolated activity began after 1.5 min. in Group C, and after 2 min. in all E groups (instructions were longer for the E groups). The interpolated activity was terminated 20 min. after the end of original learning for all four groups. Occasionally, *S*s in the E groups finished rating the passages before 20 min. had elapsed. In this case, *S* was asked to reread and rate the same passages in terms of the way one of his parents would feel about taking the trip, or about *S*'s taking the course. If less than 2 min. remained, *S* was simply asked to go back over his ratings and make sure that they accurately reflected his preferences.

An unpaced test of recall of the paired-associates list began 21 min. after the end of original learning. Each stimulus word was presented alone in the drum window, and *S* was to call out the word that had been paired with it on the list. The drum was not advanced until *S* had indicated he was ready to go on. The order of presentation of stimulus words was the order that would have occurred next if original learning had continued. Response words were not shown during the test.

A tape recording was made of the entire recall test, without *S*'s knowledge. After the experimental session was concluded, the tape was played back and a stopwatch was used to measure the latency of each correct response. The watch was started at the sound of the advance of the drum to a new stimulus word and was stopped when *S* gave the correct response. If *S* corrected himself before going on to the next stimulus word, latency was still defined as the time from the advance of the drum to the occurrence of the correct response. Because of an oversight in running the first two blocks of the experiment, tape recordings were not obtained; latency measures, therefore, were not available for the first two *S*s in each group.

The experiment was run in 16 blocks of four, with each group represented once in each block. The assignment of starting orders in original learning and of Set A or Set B passages in interpolated learning was balanced within groups. In every 2 blocks of the experiment, Set A and Set B occurred once in each group; assignment of sets was made in the control as well as experimental groups for the purpose of data analysis. The order of running groups within blocks was random. The *S*s were 68 undergraduates, men and women at the University of California, Irvine, who were fulfilling a course requirement. The *S*s were not necessarily naive to verbal-learning experiments. The *S*s were assigned to groups in order of their appearance at the laboratory. Of the 68 *S*s 4 were replaced because of experimenter error or interruptions from outside sources.

RESULTS AND DISCUSSION

Original learning. The mean number of trials to the criterion of 7/10 correct reponses on the list of paired associates was 5.2, 5.6, and 5.0 for Groups C, E–R, and E–RS_1, respectively. A two-way analysis of variance (groups; Set A or B in interpolated learning) yielded no significant *F*s (all $ps > .25$). The mean number of trials to criterion in Group E–RS_2 (4:5-sec. rate of presentation) was 2.2, and the difference between E–RS_1 and E–RS_2 was significant [$F\ (1, 28) = 10.0, p < .01$].

Interpolated reading. It is of interest to note the difference in the number of passages completed by Group E–R, in which subjects only read and rated the passages, and Groups E–RS_1 and E–RS_2, in which subjects also looked for words from the paired-associate list while making their ratings. Almost all subjects (15 out of 16) in Group E–R had progressed to the second rating task (parent's preferences) before interpolated reading was terminated; of the 32 subjects in Groups E–RS_1 and E–RS_2, only 1 had gone this far by the end of interpolated reading.

For each subject in Groups E–RS$_1$ and E–RS$_2$, the score for the discovery of words from the paired-associate list was expressed as a proportion: total number of paired-associate words underlined / total number of paired-associate words encountered. The mean proportion of list words underlined was high: .81 and .77 in Groups E–RS$_1$ and E–RS$_2$ respectively. There were relatively few cases in which a nonlist word was underlined (a total of 20 instances in Groups E–RS$_1$ and E–RS$_2$ combined). It is clear that the task of searching for paired-associate words was fairly accurately accomplished by subjects in both groups, and that this additional task reduced the amount of new material which was read during the retention interval, relative to Group E–R. It may be the case, of course, that the subjects in Groups E–RS$_1$ and E–RS$_2$ read the same *amount* of material as did the subjects in Group E–R but that much of this reading was rereading of the same material while searching for words from the list of paired associates.

Recall of paired-associates list. A within-subjects comparison of the recall of pairs with *words present* in interpolated reading versus pairs with *words absent* in interpolated reading, was permitted by the balanced assignment of subjects in experimental and control groups to either Set A or Set B in interpolated reading. The assignment of subjects to sets in Group C was, of course, a pseudovariable and was used only for the purpose of data analysis. Group C should show no difference in the recall of words-present and words-absent pairs.

As to accuracy of recall, the mean number of words correctly recalled from the paired-associate list was essentially the same in all groups: 8.6, 8.6, 8.4, and 8.8 for Groups C, E–R, E–RS$_1$, and E–RS$_2$, respectively [F (1, 56) $<$ 1.00]. Similarly, a three-way analysis of variance for the experimental groups alone (groups; words present versus words absent; Set A or B in interpolated reading) yielded no significant main effects or interactions (all ps $>$ .10). It is clear that in terms of number recalled, interpolated reading did not lead to contexual retroactive inhibition (control versus experimental groups), nor was there specific retention decrement for those pairs which appeared in interpolated reading (words present versus words absent in the experimental groups).

With respect to response latency, each subject's mean latency was obtained for (*a*) all correct responses at recall, (*b*) the correct responses for word-present pairs, and (*c*) the correct responses for word-absent pairs. A two-way analysis of variance of mean latency (groups; Set A or B in interpolated reading) for all responses correctly recalled yielded no significant main effects or interactions (all *ps* $>$ .10). The mean latency for word-present and word-absent pairs is shown for each group in Table I. As expected, the control group showed no difference in mean latency for word-present and word-absent pairs. In the three-way analysis of variance for the three experimental groups alone, the difference between word-present and word-absent pairs was significant [F (1, 36) $= 8.13$, $p < .01$]. Neither the effect of groups nor the interaction of groups and words present versus words absent was significant (Fs < 1.00).

TABLE I

MEAN LATENCY OF WORD-PRESENT AND WORD-ABSENT RESPONSES AT RECALL

	Group			
Type of pair	C	E-R	E-RS$_1$	E-RS$_2$
Word absent	2.7	3.0	3.0	3.0
Word present	2.7	4.1	4.9	3.8

Note: Latencies in sec.

In sum, the latency of recall of the word-present and word-absent pairs was identical in the control group, while for all three experimental groups, the latency of recall was significantly greater for the word-present than for the word-absent pairs. The results indicate that the appearance of words from the list of paired associates during interpolated reading leads to interference with the recall of the corresponding pairs from the paired-associate list. The failure to find an interaction of groups and words present versus words absent in the experimental-group analysis suggests that neither knowledge of occurrence of words from the paired-associate list during interpolated reading, nor searching for the words, nor rate of list learning has differential effects on retention decrement.

A previous study demonstrated interfering effects of interpolated learning on first-list recall and on response latency with lists that conformed to the A–B, A–C, A–D paradigm; the results for both accuracy and latency of recall yielded significant retroactive inhibition.[5] The increase in latency for the experimental groups in that study was found for a set of pairs on which there was no difference between control and experimental groups in number recalled ('strong' pairs as defined by scores in original learning), as well as for the entire set of responses recalled in control and experimental groups. It may be assumed either that response latency is a sensitive measure of the same process that underlies complete failure of recall, or that two different processes are involved. An increase in response latency might reflect competition, a temporary blocking of the correct response, while failure to recall a response on an unpaced test reflects unlearning, a reduction in the strength of the A–B association. In the present study, retroactive inhibition was found only for response latency. The number of responses recalled in the control and experimental groups was essentially the same.

The contribution of competition and unlearning to response latency cannot be differentiated on an empirical basis but is clearly in the domain of theory. The present writer assumes that there is a continuum from slow responding to failure of responding and that the same basic process underlies both changes. It could thereby be expected that interpolated reading, if extended or modified in terms of frequency of occurrence of list words, would lead to a failure of recall as well as an increase in latency of recall of responses from the paired-associate list.

Regardless of the locus of interference, the results clearly demonstrate that reading—an activity which is quite different from the rote learning of a paired-associate list—leads to a significant decrement in performance on the recall test. Furthermore, the decrement is *specific* to pairs whose members appeared in the interpolated material. Considering the relatively short duration of interpolated reading in this experiment, and the difference in the nature of original learning and the interpolated activity, the occurrence of any retroactive inhibition at all is perhaps surprising.[6] The results certainly encourage a further investigation of extraexperimental interference from subsequent language activity. The forgetting that is observed when subjects have learned a single list may indeed depend in part upon interference from subsequent verbal activity outside of the laboratory.

---

[5] I. M. Birnbaum, Unlearning in two directions, *J. exp. Psychol.*, 72, 1966, 61–67.

[6] Postman (n. 1).

ERRORS IN TACTILE LOCALIZATION

CHARLES M. CULVER, *Dartmouth Medical School*

Abstract. In a test of tactile localization, subjects were lightly stimulated at points on their palm and then indicated on outline drawings where they felt they had been stimulated. Both the amount and direction of localization error were measured. There was a slight tendency for all subjects to displace their judgments toward their thumbs. Left-handed subjects displaced their judgments in a wristward direction significantly more than did right-handed subjects. The overall amount of localization error was not related to hand or eye dominance but was significantly related to a measure of visual suppression.

A test of tactile localization was developed by the author as one of a battery of instruments designed to measure spatial-perceptual abilities. The test, described in detail below, involves stimulating a subject with a von Frey hair at a series of points on the surface of his palm, which is hidden from his sight; after each stimulation the subject is asked to make a mark on an outline drawing of his hand at the point where he feels he has been stimulated. There is almost always some error on each trial, and both the extent of this error and the direction in which it occurs may be considered.

It became apparent with repeated usage of the test not only that subjects differed in their general ability to localize points but also that certain constant directional errors occurred. The author is also interested in laterality phenomena and noted that right-handed and left-handed individuals seemed to make somewhat different directional errors: left-handers frequently displaced their judgment toward the wrist. The data reported below were collected in order to examine more fully the relationship between tactile-localization performance and laterality in a group of subjects for whom relatively satisfactory measures of hand and eye dominance existed.

Received for publication by Professor E. B. Newman. The research reported was conducted at the Duke University School of Nursing and was supported, in part, by the United States Public Health Service (Research Grant NU-00085) and by a Duke University Research Council grant. The author thanks Pamela Dixon for her help with collection of the data.

METHOD

Subjects. The *S*s were 75 young women, ages 17 to 18, who made up an entire freshman class at the Duke University School of Nursing. They cooperated as a class project in return for a reimbursement to their class treasury.

Tactile localization. During the testing *S* sat across a small table from *E*. A curtain was stretched across the table between the two so that only eye contact could be made. On a given series of trials *S* extended the hand being tested underneath the curtain toward *E*. On her side of the table, *S* was provided with an outline drawing that resembled her stimulated hand; with her nonstimulated hand she held a pencil and after each trial made an X on the outline drawing at the point she felt had just been stimulated. *S* numbered these Xs sequentially to avoid any confusion in scoring.

Each *S* was stimulated at the same nine points on each palm. These 18 points were given in the following four-block sequence: points 1–5 of the nondominant hand (so that the initial Xs could be marked with the usual writing hand); points 6–9 of the dominant hand: points 6–9 of the nondominant hand; and points 1–5 of the dominant hand. At the beginning of each block of trials, *S* was provided with a fresh form of the hand outline.

The nine points stimulated on each palm were located as shown in Figure 1. Stimulation was made with a transient pressure of a von Frey hair (a nylon monofilament embedded in a Lucite handle) calibrated to deliver 3.84 g. at the point of initial bending. This stimulus strength is nearly always above palmar threshold; on the few occasions when *S* claimed not to have felt the stimulus, it was repeated.

Each *S*'s errors were scored in the following way. The points actually stimulated were superimposed on each *S*'s outline drawings. Each 'true' stimulated point was considered as the origin (0, 0) of a two-dimensional coordinate system (x, y); the *S*'s estimated point was then located in this system by being assigned x and y values in millimeters (e.g., $-3, +4$). Constant directional errors could then be assessed for each palm by summing x values and y values separately across points. It was also possible to determine the direct distance between the stimulated and the estimated points (herein called the gross error) by solving the Pythagorean theorem (e.g., $-3, +4$ equals a gross error of 5 mm.).

Hand dominance. Hand dominance was defined as that hand with which *S* wrote.

Eye dominance. The *S*s were given two tests (glasses were not worn during either test). In the test of dominance during near sighting, *S*, seated approximately 36 in. from *E*, fixated on *E*'s open right eye (*E*'s left eye was covered) with both of her eyes. *S* then raised a pointer midway between *E* and *S* so that it was 'lined up' with *E*'s right eye. This alignment was made three times with the pointer held in *S*'s dominant hand and three times in her nondominant hand. Dominance was said to exist when at least five of the six alignments were made with one eye. In the test of dominance during far sighting, *S* stood 15 ft. from *E* and held in front of her with both hands a 2 × 3 ft. piece of cardboard containing in its center a hole that was 1 in. in diameter. *S*

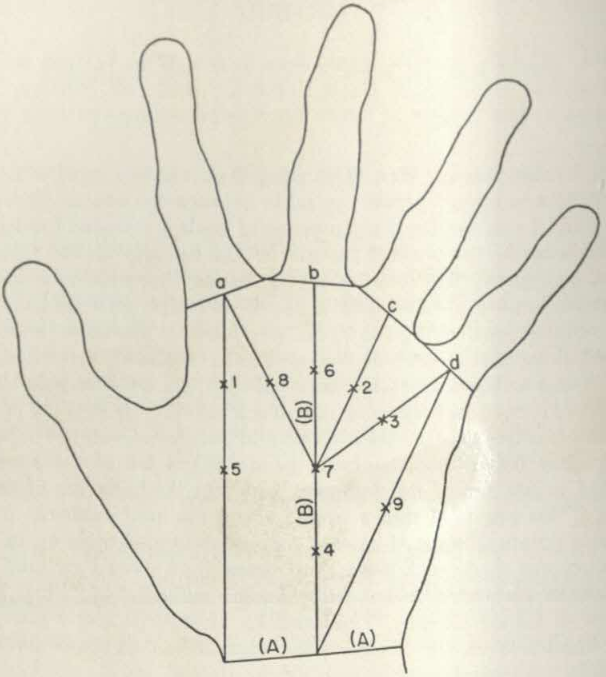


FIG. 1. Stimulated points, as determined by the following method: line A (superimposed on distal bracelet crease) was drawn and bisected; line B was drawn and bisected; all other lines were then drawn. Points a, b, c, and d represent the midpoints of proximal metacarpophalangeal creases. All stimulated points are indicated by numbers and are at bisection points of lines. Point 1 is midway between points a and 5.

bifixated on *E*'s open right eye, then raised the piece of cardboard at arms' length until she could see *E*'s open eye through the hole. Dominance was said to exist when at least three of the four sightings were made with one eye.

Of 66 right-handed *Ss*, 34 were right-eyed by both of these tests, 25 were left-eyed, and 7 were not consistently lateralized. Of 9 left-handed *Ss*, 5 were right-eyed, 2 left-eyed, and 2 not consistently lateralized.

RESULTS

The average gross and directional errors for right-handed and left-handed subjects for each of the 18 palmar points are shown in Figure 2. Table I gives numerical summaries of these data and group mean values.

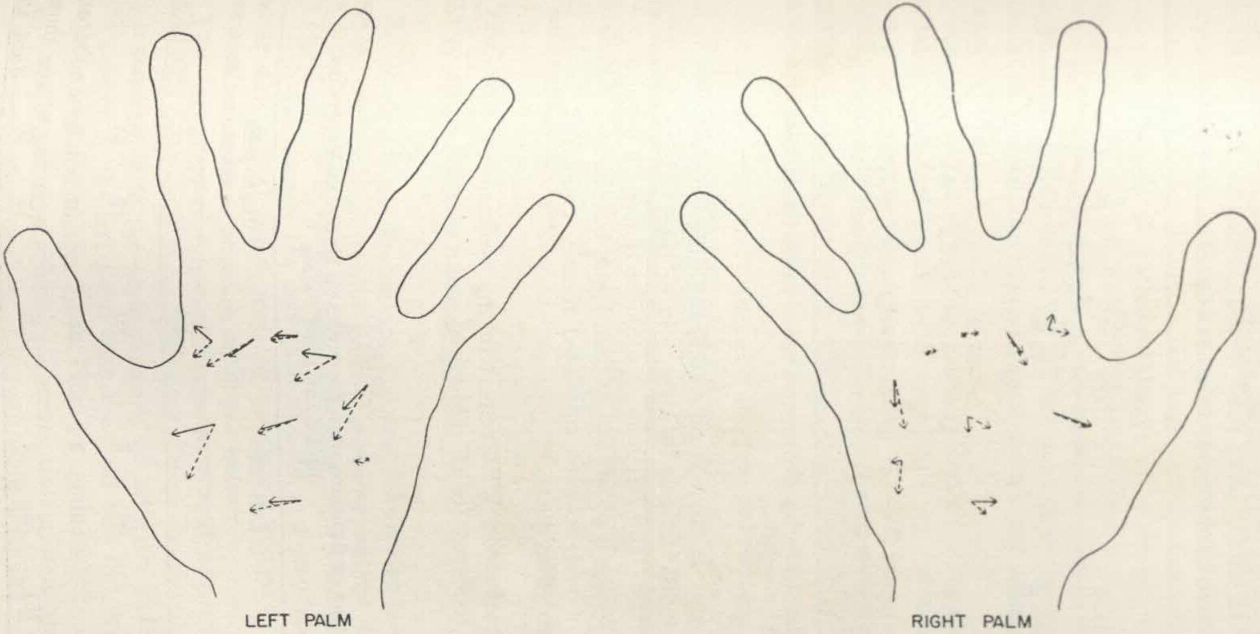


Fig. 2. Average directional errors for right-handed subjects (solid line) and left-handed subjects (broken line), for each of the 18 stimulated palmar points. The origin of each vector represents the point stimulated; the termination represents the group mean error.

TABLE I

PALMAR DIRECTIONAL-ERROR AND GROSS-ERROR VALUES

Values	Right-handed subjects $(N = 66)$	Left-handed subjects $(N = 9)$	t ratio
Total x	−47.1	−40.7	.29
Left-palm x	−73.6	−87.3	.88
Right-palm x	+26.5	+46.6	1.43
Total y	−37.9	−114.5	2.00*
Left-palm y	−14.4	−65.4	2.29*
Right-palm y	−23.5	−49.1	1.14
Total gross error	279.0	313.4	1.39
Left-palm gross error	145.5	174.0	1.85
Right-palm gross error	133.5	139.4	.46

Note: Values for each palm represent sum of values for nine points stimulated.
* $p < .05$, two-tailed.

Gross error. There was no significant difference between right- and left-handed subjects for right, left, or total palmar gross error. Both the right- and left-handed subjects tended to perform more poorly on their left palm than on their right, but this was significant ($p < .05$, t test between correlated means, two-tailed) only for the left-handed subjects.

Directional error. With respect to the x-dimension, there was a consistent tendency for all subjects to make constant errors toward

TABLE II

NUMBER OF SUBJECTS WITH POSITIVE AND NEGATIVE VALUES ON RIGHT AND LEFT PALMS

Group	Left-palm x values		p value	Right-palm x values		p value[†]
	−	+		−	+	
Right-handed $(N = 66)$	64	2	.001	15	51	.001
Left-handed $(N = 9)$	9	0	.01	0	9	.01
Total $(N = 75)$	73	2	.001	15	60	.001

[†] Sign test, two-tailed.

the thumbward, or radial, side of each palm. Table II shows the number of subjects within each group whose total left and right palmar x values were negative and positive (negative left and positive right palmar x values represent thumbward deviations). Both the right- and left-handed subjects showed a significant incidence of thumbward deviations as evaluated by sign tests. This tendency was more marked for the left palm than the right, as can also be observed in the mean left and mean right palmar x values shown in Table I.

While there was a general tendency for all subjects to make constant directional errors in the negative y, or wristward, direction, this was more pronounced among the left-handed subjects. Table I shows that the left-handed subjects had a significantly more negative total y deviation ($p < .05$); this was especially marked on the left palm ($p < .02$).

Eye dominance. The relationship between handedness, eyedness, and localization errors was explored for those subjects who showed a consistent ocular lateralization on both tests of eye dominance. No significant differences in gross error, x, or y deviations were found between right-eyed or left-eyed subjects overall, or between right-eyed and left-eyed subjects within either the right-handed or left-handed subgroups.

DISCUSSION

Relatively little is known about factors influencing tactile-localization performance. While grossly impaired ability to localize points on the body surface has some value as a localizing sign of parietal-lobe disease, the correlates of good and poor performance within the normal range and the nature of constant directional errors have been only scantily explored.

Two dimensions of directional error were analyzed separately in these data and one of these, deviations toward the thumb, was seen with consistency in the great majority of subjects. This tendency was greater on the left palm than the right and did not relate to hand or eye dominance. Both Pillsbury and Parrish also found some nearly universal directional errors after stimulating subjects on the inner surface of their forearms: points near the wrist tended to be displaced toward the wrist and points near the elbow toward the elbow.[1] It might be supposed from our data that displacement occurs toward regions of denser sensory innervation or greater area on the parietal sensory homunculus but this is not a compelling explanation for elbow and wrist displacements. Perhaps Boring's hypothesis remains the most satisfactory for these consistently seen displacements: the subject visualizes the unseen region in terms of certain prominent anchoring points (thumb, elbow, etc.) that provide a frame of reference, and subsequently migrates his judgment toward these points.[2]

The error in the y-dimension was more complex. The subjects generally tended to err in the negative, wristward, direction, but the left-handed subjects had a significantly more negative y error than did the right-handed subjects. This finding is consistent with the author's earlier observation of increased y negativity in a different group of left-handed subjects; its meaning seems obscure, but it would be interesting to determine whether analogous differences between right-handed and left-handed subjects in direction of localizing error occur on other parts of the body surface.

After the present data were collected, the author became aware that one postulate about the physiological underpinnings of the localization process in any sense organ has been put forth by Békésy.[3] The latter writer believes that central and peripheral inhibitory processes operate so that stimuli on the receptor surface are often 'funneled' or 'displaced' and that their subjectively felt location and quality are thereby altered from physical veridicality. The present author has described elsewhere the relationship between a measure of visual suppression and subjects' poor performance on a visual spatial-perceptual task, the rod-and-frame test of Witkin.[4] Since visual suppression could be thought of as a form of sensory inhibition, it seemed logical to inquire whether any relationship might have existed between the extent of the present subjects' visual suppression and their tactile-localization performance.

One measure of visual suppression was available. At the time of testing for dominance during near sighting, the subjects had been graded (for other purposes) on the extent to which they saw 'double images.' Since during this testing the subject was bifixating a far point and interposing a 'near' pointer for alignment purposes, the pointer could potentially be seen doubly, since its images in the two eyes fell on noncorresponding retinal points. Subjects vary in the extent to which they perceive double images in this situation, and the author has developed an ordinal scale ranging from 0 (double images never seen) to 8 (double images always seen) for classificatory purposes. The data here were examined in two ways. First, those subjects who always saw double images ($N = 32$) had a mean palmar gross-error score of 262.9 ($\sigma = 63.6$), while subjects who never or rarely saw double images ($N = 15$) had a mean score of 324.5 ($\sigma = 86.2$); this difference was significant ($p < .01$, t test, two-tailed). Alternatively, that 15% of all subjects who performed most accurately in tactile localization ($N = 11$) had a mean double-image score of 3.7 ($\sigma = 3.2$), while that 15% who were least accurate ($N = 11$) had a mean double-image score of 6.6 ($\sigma = 2.7$); these two sets of scores, having dissimilar distributions, were evaluated by the Mann-Whitney U test and found to be significantly different ($p < .05$, two-tailed).

---

[1] W. B. Pillsbury, Some questions of cutaneous sensibility, this JOURNAL, 7, 1895, 42-57; C. S. Parrish, Localization of cutaneous impressions by arm movement without pressure upon the skin, this JOURNAL, 8, 1897, 250-267.

[2] E. G. Boring, *Sensation and Perception in the History of Experimental Psychology*, 1942, 486.

[3] G. von Békésy, *Sensory Inhibition*, 1967, 41-63.

[4] C. M. Culver, Anisometropia and the rod-and-frame task, *Percept. mot. Skills*, 25, 1967, 377-384.

These data suggest that subjects who are inclined to suppress visual double images also perform poorly on a test of tactile localization. This relationship is similar to that found in the study cited above (see n. 4), where poor performance on the rod-and-frame test was associated with a somewhat different type of visual suppression than that measured here. Taken together, the two studies raise the question of whether the nature of the sensory inhibitory processes employed by a subject may not be an important variable in understanding individual differences in perception and cognition.

TAXONOMIC INSTANCES AS RIGHT OR WRONG ITEMS AND SELECTION STRATEGIES IN VERBAL-DISCRIMINATION LEARNING

DONALD H. KAUSLER, JOAN T. ERBER, and GAYLE A. OLSON,
Saint Louis University

Abstract. Instances of a single taxonomic category served as either the right (R) or wrong (W) items of a verbal-discrimination list. Relative to a control condition having unrelated items on the list, learning was markedly accelerated when the instances were R items, but it was only slightly accelerated when the instances were W items. Additional evidence provided by the free recall of R and W items supported the conclusion that a general selection strategy mediated the verbal-discrimination learning only when the positive instances defining class membership occurred as R items.

According to the frequency theory of verbal-discrimination learning, intrapair discriminations between wrong (W) and right (R) items are based on the rote accrual, during practice, of a greater response frequency to R items than to W items, and on the subsequent selection of intrapair items with the greater frequency (Rule 1 of frequency theory).[1] However, the rote accrual of response frequencies as a prerequisite for intrapair discriminations may be largely abrogated by the activation of a general selection strategy. A selection strategy is likely to be activated whenever the list is structured such that either the R items or the W items are members of a discriminable class. Selection of R items may then be made on the basis of class membership alone (i.e., choosing or avoiding class instances, contingent upon the W–R locus of the class). The efficacy of a selection strategy in mediating verbal-discrimination learning was demonstrated previously, by Kausler and Farzanegan, for lists in which preexperimental word frequency provides the attribute for class membership.[2] The present study tested further the efficacy of a

---

Received for publication April 22, 1970. The authors wish to thank Sylvia Kollasch for her assistance in data collection and analysis.

[1] B. R. Ekstrand, W. P. Wallace, and B. J. Underwood, A frequency theory of verbal-discrimination learning, *Psychol. Rev.*, 73, 1966, 566–578.

[2] D. H. Kausler and F. Farzanegan, Word frequency and selection strategies in verbal-discrimination learning, *J. verb. Learn. verb. Behav.*, 8, 1969, 196–201.

selection strategy when a taxonomic category provides the attribute for class membership.

Instances of a taxonomic category (e.g., *animals*) are likely to be associatively related, either directly (e.g., *dog* ⇄ *cat*), or indirectly by mediation via the common superordinate (e.g., *dog* ⇄ [*animal*] ⇄ *bear*). Consequently, the effect of a categorically based selection strategy on intrapair discriminations may be modified by a source of rote frequency units—implicit associative responses[3]—that does not need to be considered when associatively unrelated words constitute the class membership. If items in a verbal-discrimination list are associates of one another, then responses to a given item add frequency units both to that item and to its associates in the list, the former directly and the latter indirectly via the implicit associative responses emitted to the given item. Ekstrand, Wallace, and Underwood (n. 1) have offered convincing evidence for a facilitative effect on verbal-discrimination learning when the associatively related items are R components and a detrimental effect when they were W components. The increment in frequency units to R items contributed by the implicit associative responses increases intrapair W–R differentials in frequency, making Rule 1 application more efficacious. On the other hand, the increment to W items decreases intrapair differentials in frequency, making Rule 1 application less efficacious. Thus, on the basis of rote processes alone, frequency theory predicts that a list containing R items that are all instances of the same taxonomic category should be learned more rapidly than a control list containing unrelated items, whereas a list containing comparable W items should be learned more slowly than the control list. Superior performance for the R-item condition relative to the control condition therefore would not necessarily be attributable to a selection strategy (i.e., choosing intrapair components that are class members). However, performance for the W condition that is both superior to the control condition and comparable to the R condition would most likely reflect the operation of a selection strategy (i.e., avoiding intrapair components that are class members) that compensates for the detrimental effect of rote frequency processes.

As a means of gaining additional insights into the relative roles played by rote and higher-order processes during verbal-discrimination learning when taxonomic instances serve as W or R items, a free recall of all items followed verbal-discrimination acquisition for each condition in the present study. Previous research has indicated the value of free-recall measures in analyzing processes operative during the preceding verbal-discrimination acquisition.[4] Two kinds of clustering during recall were of special interest in this study. The first is the successive recall of intrapair W and R items from the verbal-discrimination list; the second is the successive recall of items that were homogeneous with respect to function on the list (i.e., functioning as W items or functioning as R items). Relative to clustering for the control condition receiving unrelated items on the list, the first kind of clustering should be markedly reduced, and the second kind of clustering markedly increased, when verbal-discrimination acquisition is mediated by a general selection strategy.

---

[3] W. A. Bousfield, G. A. Whitmarsh, and J. J. Danick, Partial response identities in verbal generalization, *Psychol. Rep.*, 4, 1958, 703–713.

METHOD

Design and subjects. The 60 undergraduate *S*s, all of whom were naive with respect to prior participation in verbal-learning experiments, were assigned alternately to three groups (each $N = 20$): two experimental (Groups E1 and E2) and one control (Group C). Group E1 learned a 12-pair verbal-discrimination list in which all 12 of the R items were instances of the same taxonomic category and each of the 12 W items was an instance of a different category. For Group E2 all 12 of the W items were instances of the same taxonomic category and each of the 12 R items was an instance of a different category. For Group C all 12 W items and all 12 R items were instances of different categories (i.e., the items were from 24 different categories).

Lists. For purposes of greater generalization of results and for greater control of idiosyncratic characteristics related to taxonomic categories, several different lists varying in specific item content were employed for each experimental condition and for the control condition. That is, the specific lists for a given condition varied either in the categories represented or in the instances selected from within a given category. The norms of Battig and Montague provided the source material for categories and instances.[5] Briefly, the principles followed in constructing list variations both within and between conditions were (*a*) that each category function equally often as a source of

---

[4] F. E. Fulkerson and D. H. Kausler, Effects of intrapair and interpair bidirectional associates on verbal-discrimination learning, *J. verb. Learn. verb. Behav.*, 8, 1969, 307–310.

[5] W. F. Battig and W. E. Montague, Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms, *J. exp. Psychol. Monogr.*, 80, 1969, (No. 3).

W and R items, (*b*) that items from a given category always be selected from the first 20 responses in the normative data, the lists being made as comparable as possible with respect to the mean ranking of items within their respective categorical response hierarchies, (*c*) that lists be as comparable as possible with respect to such item characteristics as word frequency and graphic features, and (*d*) that associative relationships, as judged by the *E*s, be minimized as much as possible between W items for Group E1, between R items for Group E2, between both W and R items for Group C, and between intrapair W–R items for all groups. Subsequent analyses revealed that collapsing the various list forms together within each of the three main conditions did not markedly alter the outcome of tests for between-conditions effects. Consequently, for simplicity of presentation, only the results for the collapsed list forms are presented below.

Procedure. Each *S* was given standard instructions concerning verbal-discrimination learning prior to receiving one trial on a warm-up list that contained two pairs comparable to those of the experimental or control list he would be receiving. This practice list was presented in order to acquaint *S* with the nature of the task and to ensure that he understood the instructions. All lists were presented on a Lafayette memory drum at a 2:2 sec. rate, with a 4-sec. intertrial interval. First the pair appeared in the window for 2 sec., and then the same pair appeared in the window for another 2 sec., this time with the correct word underlined in order to provide informative feedback. All *S*s were instructed to respond to each pair on each trial of the learning task, including the first. Verbal-discrimination learning under each of the three conditions was carried to a criterion of three successive perfect trials. After reaching criterion, *S* was given a sheet of paper with 24 lines and was instructed to write down as many of the items, both R and W, as he could remember, in any order he wished, with only one item per line. There was no time limit for this task. Instructions concerning possible recall tasks were not given until verbal-discrimination learning was complete.

RESULTS

The means and standard deviations for trials to criterion on the verbal-discrimination list were 6.20 and 2.50, 8.75 and 3.01, and 10.70 and 3.97 for Groups E1, E2, and C, respectively. An analysis of variance revealed a significant between-group effect [$F\ (2, 57) = 9.82$, $p < .001$]. A Duncan's test yielded a significant contrast between Group E1 and Group C ($p < .01$), and between Group E1 and Group E2 ($p < .05$). In addition, the contrast between Group E2 and Group C approached significance ($p < .10$). The means and standard deviations for total errors in attaining the criterion were 11.00 and 8.71, 16.70 and 9.71, and 19.00 and 10.73 for Groups E1, E2, and C, respectively. The between-group effect was again significant [$F\ (2, 57) = 3.57$, $p < .05$], as was the contrast between

Group E1 and Group C ($p < .05$). The contrast between Group E1 and Group E2 approached significance ($p < .10$), but the contrast between Group E2 and Group C fell far short of significance ($p > .10$).

Item recall was scored separately for R items and for W items. For R-item recall the means and standard deviations were 9.00 and 1.45, 8.30 and 2.08, and 9.15 and 1.09 for Groups E1, E2, and C, respectively. Comparable values for W-item recall were 4.70 and 2.18, 6.50 and 1.36, and 7.05 and 2.48. The fact that the three groups differed significantly in the number of trials on the verbal-discrimination list and the fact that recall of items, may, in turn, covary with number of trials force a cautious interpretation of between-group effects on recall. However, within-group correlation coefficients between number of trials to criterion and the amount of both R-item and W-item recall were all negligible and clearly nonsignificant (all *ps* $> .10$). It seems unlikely, therefore, that between-group effects on free recall were affected by the differential number of trials.

A 3 × 2 mixed analysis of variance, with list condition as the three-level between-group variable and directionality of recall (R versus W) as the within-group variable, revealed significant main effects for groups (i.e., list condition) [F (2, 57) $= 3.84$, $p < .05$] and directionality [F (1, 57) $= 83.01$, $p < .001$], and a significant interaction effect [F (2, 57) $= 6.90$, $p < .01$]. Because of the significant interaction effect, the difference in recall attributable to list condition was examined separately for R items and for W items. The between-group effect was significant for W items [F (2, 57) $=$ 7.13, $p < .001$] but not for R items [F (2, 57) $= 1.62$, $p > .10$]. In addition, Duncan's test yielded significant contrasts in W-item recall between Group E1 and Group C, and between Group E1 and Group E2 (*ps* $< .01$), but not between Group E2 and Group C ($p > .10$).

In view of the difference in free recall between groups, between-group effects for both intrapair (W-R consecutive recall) and intracategory (consecutive-W and consecutive-R recall) clustering were analyzed in terms of relative clustering scores. For intrapair clustering the formula used was $Cl/(N/2)$, where Cl is the number of intrapair clusters during recall and N is the total number of W and R items recalled. The means and standard deviations for this score were .21 and .21, .42 and .21, and .63 and .27 for Groups E1, and E2, and C, respectively. The between-group effect was significant [χ^2 (2) = 17.20, $p < .001$], as were the contrasts between Group E1 and Group C [χ^2 (1) = 12.10, $p < .001$], and between Group E1 and Group E2 [χ^2 (1) = 4.90 $p < .05$]. The contrast between Groups E2 and C, however, did not reach significance [χ^2 (1) = 2.50, $p > .10$].

For intracategory clustering the score computed for each *S* was the difference between the observed and the expected number of intracategory repetitions, using Formula II of Bousfield and Bousfield.[6] The means and standard deviations for this score were 1.14 and 2.42, −1.39 and 1.95, and −3.14 and 2.82 for Groups E1, E2, and C, respectively. In each case, the amount of clustering was significantly greater than chance expectancy [correlated *ts* (18) = 2.11, 3.16, and 4.48, and *ps* < .05, 01, and .01, respectively]. The negative clustering scores for Groups E2 and C indicate that very little consecutive recall of either W items or R items occurred. That is, recall in these groups, as reflected in their intrapair clustering scores, consisted largely of alternating W and R items, many of which had been paired together in the original verbal-discrimination list. The between-group effect was again significant [χ^2 (2) = 14.80, $p < .001$], as were the contrasts between Group E1 and Group C [χ^2 (1) = 8.10, $p < .01$], and between Group E1 and Group E2 [χ^2 (1) = 4.90, $p < .05$]. Again, the contrast between Groups E2 and C did not approach significance [χ^2 (1) < 1].

DISCUSSION

The present results indicate that verbal-discrimination learning is considerably more rapid for a list in which the R items are instances of a single taxonomic category than for a list in which the W items are the instances or for a list composed of completely unrelated items. As noted earlier, the rapid acquisition of an R-based taxonomic list may be the result either of a higher-order selection strategy or of a rote accrual of implicit associative responses to R items (which accrual in turn enhances Rule 1 application). Indirect support for the selection-strategy alternative stems from the free-recall data of Group E1. Recall of W items was significantly below the level of the other two groups, as is to be expected if intrapair selections of R items are mediated by a process that enables the subject to restrict his responses largely to the R items. More important, however, is the fact that clustering of items was very different from that shown in the control group. The dominant type of clustering for Group E1 was in terms of consecutive recall of items having similar functions (R or W) during original acquisition, a type of clustering that was nonexistent for Group C. A likely explanation for the disparity in clustering patterns is that Group E1's retrieval of items for recall was organized in terms of the two broad categories, member-nonmember of the taxonomy, that had served earlier to discriminate between paired items.

---

[6] A. K. Bousfield and W. A. Bousfield, Measurement of clustering and of sequential constancies in repeated free recall, *Psychol. Rep.*, 19, 1966, 935–942.

On the other hand, the present results indicate that a general selection strategy is unlikely to be operative when the W items are the instances of the taxonomic category. Militating against the notion that Group E2's verbal-discrimination acquisition was mediated by a selection strategy is the close correspondence between free recall for this group and free recall for Group C. Especially convincing is the failure to find clustering for Group E2 on the basis of item functioning. It is conceivable that selection strategies in general are unlikely to be activated when the positive instances of the category occur only in W items. The nature of the verbal-discrimination task is such that the subject's activity during informative feedback phases of practice is ordinarily directed only to R items. However, it has been demonstrated that forced pronunciation of W items during feedback phases alters the learning processes,[7] and it is likely that a similar procedural requirement when W items are positive instances would succeed in activating a selection strategy.

The prediction that Group E2, in the absence of a selection strategy, would learn the verbal-discrimination list more slowly than Group C was not fulfilled. In fact, Group E2 displayed slightly faster learning than Group C. This finding of faster verbal-discrimination learning when implicit associative responses are elicited to W items than when items are unrelated is in agreement with the results of a recent study by Kanak, Cole, and Thornton.[8] The most likely explanation for this finding is that the occurrence of implicit associative responses to W items may be great enough to yield summated response frequencies to W items that exceed the summated frequencies to their paired R items. Under these circumstances, intrapair discriminations may be expedited by the use of Rule 2 of frequency theory ('always select the intrapair component with the smaller frequency'; see n. 1).

---

[7] D. H. Kausler and R. J. Sardello, Item recall in verbal-discrimination learning as related to pronunciation and degree of practice, *Psychon. Sci.*, 7, 1967, 285–286.

[8] N. J. Kanak, L. E. Cole, and J. W. Thornton, Inter-item manipulations of implicit associative responses in verbal-discrimination learning, *Psychon. Sci.*, 18, 1970, 359–361.

NOTES AND DISCUSSION

THE INTERRELATIONSHIPS AMONG A NUMBER OF ASSOCIATION MEASURES

Abstract. The five-letter words used by Taylor and Kimble were rated by 100 subjects for number of associations, familiarity, and pronounceableness; and Glaze's measure was derived from their responses to the first measure. The correlations among these measures of association were quite high (.90–.98), and the relationships between them and the Taylor-Kimble measure of association latency were also substantial (.73–.78).

Taylor and Kimble used a variant of Glaze's procedure to provide a meaningfulness, or association, measure for 319 five-letter, structurally homogeneous (CVCVC) nouns and paralogs.[1] Their procedure consisted of presenting each verbal item to each subject with instructions to give the first association to that item as quickly as possible. A variety of response measures was obtained; the most conventional and Glaze-related of these was the percentage of subjects responding to each item in less than 2.5 sec.

The population of verbal items utilized by Taylor and Kimble (n. 1) is an excellent source of items for investigators desiring to use material of a somewhat more complex nature than the traditionally used trigram. It would be of value to have other association measures for these items, and to be aware of the relationship between these measures and the association measure used by Taylor and Kimble. The basic purpose of this study was to obtain measures of rated association value, familiarity, and pronounceableness for the population of items provided by Taylor and Kimble, as well as to examine the interrelationships among all of these measures.

Received for publication May 4, 1970.

[1] J. D. Taylor and G. A. Kimble, The association value of 320 selected words and paralogs, *J. verb. Learn. verb. Behav.*, 6, 1967, 744–752. See also J. A. Glaze, The association value of non-sense syllables, *J. genet. Psychol.*, 35, 1928, 255–269. Taylor and Kimble have indicated that there were 320 items in their population of items. A count of those listed revealed just 319.

METHOD

Materials. The stimulus materials were the 319 five-letter words and paralogs of a CVCVC structure utilized by Taylor and Kimble. Three seven-point rating scales were constructed to measure: (*a*) number of associations, (*b*) familiarity of subjective frequency, and (*c*) pronounceableness. The scale of pronounceableness used the category descriptions of 'very hard,' 'hard,' 'below average,' 'average,' 'above average,' 'easy,' and 'very easy.' The familiarity scale was similar to that used by Noble;[2] here *Ss* were instructed to rate the verbal items as to the number of times they had seen, heard, or used them in speaking, writing or reading. The categories used were 'never,' 'rarely,' 'below average,' 'average,' 'above average,' 'often,' and 'very often.' The association scale, modeled after that used by Noble, Stockwell, and Pryer, required *Ss* to judge the number of associations they had to each of the presented items.[3] The categories employed were 'none,' 'few,' 'below average,' 'average,' 'above average,' 'many,' and 'very many.' Finally, it was possible to use this scale to approximate Glaze's measure (n. 1) by noting the relative frequency of responses in categories beyond 'none.' In an effort to provide a task that could be readily completed within a 45-min. period, two forms of each scale were constructed, with each form consisting of 159 or 160 items. The construction of the two scales was accomplished by using the ordering of items on association value as reported by Taylor and Kimble, and assigning items with odd rank to one scale and even-ranked items to the other.

Subjects and procedure. The *Ss* were 100 introductory psychology students. Tested in small groups, they were given the three scales and asked to provide ratings for each of the CVCVCs listed.

RESULTS AND DISCUSSION

Fewer than 50 of the possible 45,850 responses were lost because of the subjects' failure to respond to a particular item. Thus at least 49 responses were available from which to compute a mean for each item on each scale. The reliability of each of the three rating scales was assessed by correlating mean ratings obtained from a randomly selected half of the subjects with corresponding ratings obtained from the other half of the subjects. The reliability coefficients for the pronounceableness, association, and familiarity scales were .94, .98, and .98, respectively.

The correlations obtained among the varying measures of association, including Taylor and Kimble's association-latency measure, can be found in Table I, where it may be noted that the interrelationships among the various scale measures were very high; the relationships between these measures and the Taylor and Kimble measure were also substantial; thus, it is likely that any one of these measures would be adequate for most investigators' use.

---

[2] C. E. Noble, The meaning-familiarity measure, *Psychol. Rev.*, 60, 1953, 89–98.

[3] C. E. Noble, F. E. Stockwell, and M. W. Pryer, Meaningfulness (m') and association value (a) in paired-associate syllable learning, *Psychol. Rep.*, 3, 1957, 441–452.

TABLE I

INTERCORRELATIONS AMONG ASSOCIATION MEASURES

	Rated pronunciation	Rated association	Rated familiarity	Glaze association
Rated association	.92			
Rated familiarity	.91	.98		
Glaze association	.91	.92	.90	
Taylor and Kimble association	.74	.78	.76	.73

These general findings are supportive of, and extend the results obtained in, an earlier study by Hall, who also found that the interrelationships among measures of rated association, familiarity, and pronounceableness for 200 randomly selected trigrams were quite high, although the correlations he got did not reach the values obtained here.[4] The findings also support a recent factor-analysis study by Frincke, who, examining a variety of word characteristics, identified what he described as a meaningful-familiarity factor, since it loaded heavily on association, familiarity, and objective-frequency measures.[5]

JOHN F. HALL, *Pennsylvania State University*
LAWRENCE MAYS, *University of Virginia*

---

[4] J. F. Hall, Relationships among a number of measures of meaningfulness, this JOURNAL, 30, 1967, 291–293.

[5] G. Frincke, Word characteristics, associative-relatedness, and the free-recall of nouns, *J. verb. Learn. verb. Behav.*, 7, 1968, 366–372.

GUESSING AND THE ORDER-OF-APPROXIMATION EFFECT

Abstract. The efficiency with which subjects can fill a missing letter from pseudowords of zero- to fourth-order approximation to English was investigated empirically. The estimate was used to evaluate how much of the order-of-approximation effect can be attributed to guessing. It appears that guessing has negligible effects.

When letter sequences of various orders of approximation to English are presented tachistoscopically, subjects identify more letters from the higher orders of approximation to English.[1] A recent experiment has shown that the order-of-approximation effect cannot be explained in terms of simple guessing (i.e., by response-bias notions).[2] Although that experiment ruled out response-bias explanations, it did not take into account all response-bias effects. Thus, while it is clear that response-bias models do not provide an adequate account of the order-of-approximation effect, it is not clear how much of the order-of-approximation effect can be attributed to response-bias effects. The present research was designed to consider this question.

According to response-bias accounts, subjects fill in unavailable material on the basis of what the context, combined with their experience with English, indicates ought to be correct. If subjects fill in material in this way, a conservative estimate of the utility of such operations would be provided by measuring the accuracy with which subjects can fill in missing material when presented with a maximum of context.

Received for publication by Professor E. B. Newman. The research reported was supported by the Defence Research Board (Grant 9401-26 to M. P. Bryden) and the National Research Council (Grant AP-318 to the author). The author thanks Mrs. S. O. Dick for testing the subjects in the experiment proper.

[1] G. A. Miller, J. S. Bruner, and L. Postman, Familiarity of letter sequences and tachistoscopic identification, *J. gen. Psychol.*, 50, 1954, 129–139.

[2] D. J. K. Mewhort, Familiarity of letter sequences, response uncertainty, and the tachistoscopic recognition experiment, *Canad. J. Psychol.*, 21, 1967, 309–321.

METHOD

Subjects. The *S*s were 32 undergraduates at the University of Waterloo. They were tested immediately following another experiment, which involved recognition of tachistoscopically presented number sequences.

Materials and procedure. Each *S* was shown 32 eight-letter pseudowords, 8 each of zero-order, first-order, second-order, and fourth-order approximation to English. These materials were taken from the lists provided by Miller et al. (n. 1). In each pseudoword, one letter was deleted and replaced with an ampersand. Within each order of approximation to English, the position of the deleted letter was counterbalanced with pseudowords according to an 8 × 8 latin square. Counterbalancing in this way divides the stimuli into eight sets; four *S*s were randomly assigned to each set of material.

The pseudowords were typed on 3 × 5 in. file cards and were presented individually in a different random order for each *S*. Each card was presented until the *S* had completed his response.

Replication. The experiment was repeated using 32 different *S*s. For the replication, however, each set of pseudowords was typed on a single sheet and the *S*s responded by writing their response beside the appropriate pseudoword.

RESULTS AND DISCUSSION

Because the experiment and its replication yielded highly similar results, they were combined. Taking the number of letters correctly supplied as a proportion of the numbers that it was possible to obtain, the probability of correctly filling in a missing letter was estimated. The estimates for the zero-, first-, second-, and fourth-order pseudowords were .023, .068, .131, and .188, respectively.

The data clearly show that the subjects did fill in a missing letter more efficiently from the higher orders of approximation to English. The remarkable feature of the present data is that the subjects filled in letters as poorly as they did. For example, suppose that one uses the hit rates estimated in this experiment to calculate the expected value for the order-of-approximation effect in a tachistoscopic recognition experiment. To illustrate, suppose that subjects are able to report four of eight letters on the basis of the information received from the tachistoscopic exposure. In this case, they must fill in four letters by guessing from the context provided by the letters they can report. Based on the binomial distribution ($p = .188$ for fourth-order material and .023 for zero-order material; see above), the expected number of correct letters in four guesses is .092 and .752 for zero- and fourth-order materials respectively. Thus, the average number of letters correct in the hypothetical recognition experiment ought to be approximately 4.09 and 4.75 respectively.

The results of this calculation of expected means differ wildly from the typical results in tachistoscopic experiments. It has been shown under a wide range of exposure durations and levels of illumination that the mean number of letters reported from the fourth-order material is approximately 1.59 times that reported from the zero-order material.[3] Thus, if subjects report a mean of approximately 4.1 letters from a set of eight-letter zero-order pseudowords, they can typically report 6.52 letters from a corresponding set of fourth-order pseudowords. In other words, the size of the order-of-approximation effect predicted from the present estimate of hit rate is only 27% of the usual order-of-approximation effect. Furthermore, because the probability of guessing two letters in four guesses ($p = .188$, $q = .812$) is approximately .10, it is clear that the probability of guessing the prerequisite 2.42 letters is close to being significantly small by the usual standards.

Not only does an estimate of the size of the order-of-approximation effect widely underestimate typical tachistoscopic results, but the estimate itself is likely to be an overestimate. First, the estimate is based on a probability obtained from an experiment in which subjects had a maximum context (i.e., seven of eight letters). In the present example, we allowed a context of only four of eight letters. Thus, it seems reasonable to assume that the efficiency of guessing in an actual recognition experiment ought to be less for the higher order of approximation than the values obtained in the present guessing experiment. Secondly, in using the expected value of the binomial distribution, we assumed that guesses were independent. This seems an unlikely assumption. In view of the fact that the probability of guessing a letter correctly is low, it is reasonable to assume that subjects generate a good deal of misleading context. Thus, if their guesses are not independent, that self-generated context ought to further reduce guessing efficiency from the estimates provided in the present experiment. In sum, taking the present evidence together with previous studies of response-bias effects, it appears that response-bias effects such as guessing neither serve as an explanation for the order-of-approximation effect nor contribute appreciably to the size of the effect.

---

[3] D. J. K. Mewhort, Sequential redundancy and letter spacing as determinants of tachistoscopic recognition, *Canad. J. Psychol.*, 20, 1966, 435–444; D. J. K. Mewhort and E. Tulving, Sequential redundancy, illumination level, and exposure duration as determinants of tachistoscopic recognition, paper presented at the meeting of the Canadian Psychological Association, Halifax, Nova Scotia, 1964; Miller et al. (n. 1 above); E. Tulving, Familiarity of letter-sequences and tachistoscopic identification, this JOURNAL, 76, 1963, 143–146.

D. J. K. Mewhort, *Queen's University, Canada*

A CONTROL FOR OCULAR TRACKING IN THE DEMONSTRATION OF POSTRETINAL VISUAL STORAGE

If a figure moves rapidly behind a stationary opaque screen and is viewed through a narrow slit in that screen, more of the figure will be seen at any instant than is actually present in the slit and the figure will be seen as compressed along the direction of its movement.[1] One explanation for this phenomenon could be that an observer's eyes tend to track the moving figure and, therefore, to 'paint' its image across the retina. By this hypothesis, the observed shrinkage would reflect imperfect (slow) tracking. The tendency to perceive instantly more than is present would be attributed to retinal afterimages. In fact, Anstis and Atkinson have shown that when observers were *required* to track moving figures viewed through a slit, the perceived extent of each stimulus pattern could be predicted from the extent of the eye movements that occurred during its presentation.[2] There is little ground for doubting the proposition that tracking, *when* it occurs, would produce percepts of the kind described above. However, Anstis and Atkinson concluded that *any* instance of figural perception under these conditions depends on such movements. A simple demonstration (Figure 1) suffices to bring that generalization into question.

---

[1] S. M. Anstis and J. Atkinson, Distortions in moving figures viewed through a stationary slit, this Journal, 80, 1967, 572–585; P. A. Kolers, Some psychological aspects of pattern recognition, in P. A. Kolers and M. Eden, *Recognizing Patterns*, 1968, 49; J. Hochberg, In the mind's eye, in R. N. Haber (ed.), *Contemporary Theory and Research in Visual Perception*, 1968, 314–316; N. H. Haber and L. S. Nathanson, Post-retinal storage? Some further observations on Parks' "Camel as seen through the eye of a needle," *Percept. & Psychophys.*, 3, 1968, 349–355; T. E. Parks, Further comments on the evidence for post-retinal storage, *Percept. & Psychophys.* (in press). See Anstis and Atkinson for an excellent review of earlier reports. An alternative to the tracking hypothesis mentioned in the text has been proposed by W. C. Hoffman, The lie algebra of visual perception, Mathematical Note 408, Boeing Scientific Research Laboratories, 1965, 47.

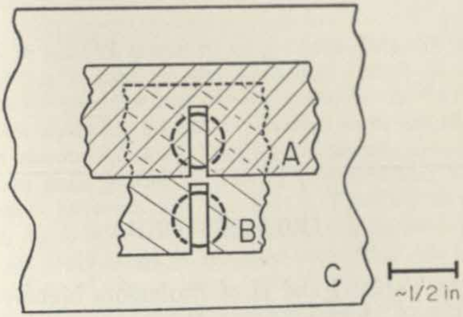


FIG. 1. Arrangement of the demonstration. Materials include three pieces of common opaque card-stock. A simple figure (a circle) is drawn on surface C and viewed through a slit in card B. An identical figure is drawn on card B and viewed through a slit in card A. Cards A and C are held stationary, while B oscillates between them at a rate of approximately one cycle per second. The vicinity of the upper edge of the lower figure forms a convenient fixation point for the simultaneous observation of both stimuli. Viewing distance should be 15 in. or more.

All that is required is a stimulus situation which allows for simultaneous observation of (*a*) a moving figure viewed through a stationary slit and (*b*) an identical, but stationary, figure viewed through a moving slit. Under these conditions, any tracking movements that are present in the observation of the first (moving) figure will also result in a perceived distortion in the shape of the second (stationary) figure. That is, Haber and Nathanson (n. 1) have shown that tracking a slit which moves over a stationary stimulus will produce shrinkage in the apparent shape of that stimulus. On the contrary, however, my observations under the conditions described in Figure 1 often include a (foreshortened) percept of the moving circle coincidental with veridical perception of the stationary circle.

---

[2] Anstis and Atkinson (n. 1); they did not, however, include a no-movement condition. That condition would, of course, be critical to the present question.

THEODORE E. PARKS, *University of California, Davis*

CORRECTIONS

The notes to Tables I and II of Professors Mershon and Gogel's article on "Effect of Stereoscopic Cues on Perceived Whiteness" (Vol. 83, No. 1, pp. 61 and 62) should both read:

> Note: On scale of W', full white is 6 and black is 1; d' and D' in in.

The seventh line of the third paragraph of Father Stafford's article on "Thomas Verner Moore: 1877–1969" (Vol. 83, No. 2, p. 286) should read:

> laboratory at Catholic University in 1891, to Leipzig to study under

BOOK REVIEWS

Edited by JULIAN HOCHBERG,
Columbia University

THE EXPERIMENTAL PSYCHOLOGY OF SENSORY BEHAVIOR. By JOHN F. CORSO. New York: Holt, Rinehart and Winston, 1967. Pp xii, 628. $12.95.

Suppose I say to you: "I intend to write a book that will prepare the unprepared for a course in experimental psychology. I shall do this by dealing mainly with sensory processes and research methods. In this way I can at once cope with the boundless growth of published work and show the impact of this work on sensory behavior. First I shall prepare the student for discovering empirical laws relating stimuli to responses and for formulating scientific theories of behavior. For this, Part I, 'Foundations of Psychology' is indispensable, as it will cover philosophy, physics and physiology–the disciplines that are the bedrock of scientific psychology. Next I shall try to remedy the inability of the student to read original papers by relating theories of measurement to the design of experiments and to psychophysical methods. This aim is to be achieved by Part II, 'Quantitative Methods in Psychological Research.' In Part III, 'Classical Problems in Experimental Psychology,' the student will be led by way of sensory thresholds considered classically, through adaptation and recovery, to disorders of the eye, ear, balance, smell, and taste. What was learned in physics and, especially, physiology often will allow these phenomena to be explained by the ideas of neurology. Both normal and abnormal sensory behavior will be explained in the context of selected theories. From the islands of Psychophysics and Sensory Specialists (lying just off the coast of Physics and Physiology) the student will be airlifted out of the classical problems to explore four selected clearings in the Jungle of Theory. The vehicle is Part IV, 'General Theories in Contemporary Psychology.' One chapter each will be devoted to (*a*) theories of sensory discrimination, (*b*) information theory, (*c*) adaptation-level theory, and (*d*) theories related to sensory deprivation. These are general, not special, theories. They are not specific to any one sense mode. In short, I shall give the student a wide view of theory and experiment as practiced by contemporary sensory and perceptual psychologists."

Those are not Professor Corso's words: they are mine. But I believe them to hold the essence of his words about the aims and arrangement of his book, and the reasons for writing it. I accept most of his aims and intentions as being valid and sensible. Anyone who has thought of writing a monograph on just one of the perceptual systems—even assuming that the student knows mathematics, physics, and physiology—must have a deep appreciation of the very ambitious task that Professor Corso set himself. Many of us had hoped that his book would provide us and our students with a contemporary review of psychophysics and sensory psychology. I am sorry to have to say that being mortal like any of us, he has fallen short of most of the goals. But he has put us in his debt, for his heroic try reveals the true vastness of the job and points out directions to be tried and byways to be shunned.

What went wrong? Said briefly, I believe that the several parts were not tied together to make a unified whole. The parts are incomplete or idiosyncratic and do not take proper account of important contemporary work in experimental and theoretical psychology. I shall give now examples of some of the most grievous faults.

Some topics suffer from an unwarranted incompleteness or omissions. Color vision is discussed almost incidentally in the 10 pages devoted to deficiencies in color vision. About theory Professor Corso says, "Our presentation, however, will be restricted to the Young-Helmholtz theory; the reader may pursue the other theories, namely, the Hering theory, the Ladd-Franklin theory, and the Land theory, in other sources" (p. 370). The fundamental notion of recurrent inhibition is not treated, and the names of Hartline, Granit, or Ratliff do not occur. Lateral inhibition and reciprocal inhibition are mentioned once in passing, and it is recognized that the former may be a general principle of afferent systems (though definitions given by example are confusing). F. Ratliff's book, *Mach Bands: A Quantitative Study on Neural Nets in the Retina*, 1965, is, astonishingly, not referenced. There are a number of psychophysical methods not mentioned. In the section on signal-detection theory there is no mention of the 2, 3, . . . , n-interval temporal or spatial forced-choice method or its relation to the d' measure of the 'yes-no' experiment.

There are idiosyncratic and misplaced emphases. The dead issue of neural quantum theory of sensory discrimination is given 23 pages but the lively theory of signal detectability only 15. The exposition of the latter is not easy to follow, in part because there are no good computational examples. Incidentally, the important book of D. M. Green and J. A. Swets (*Signal Detection Theory and Psychophysics*, 1966) is not referenced. There are 5½ pages on factors in conductive deafness, 6½ pages on factors in neural (perceptive) deafness, and 3½ pages on hearing aids and auditory theory.

It is not easy to justify such imbalances and omissions as due to selectivity. The theory of hearing that began with Békésy in about 1928 has generated a great amount of important work, ranging from molecular biology through auditory psychology and physiology to the perception of music. Yet Professor Corso has virtually omitted any reference to this work. The ideas of lateral inhibition in retinal nets, based on the work of Hartline and Granit, began over 30 years ago. These ideas, coupled with those about pattern and motion detectors in the retina, have enormously increased our understanding of the way in which general neural nets work. Experimental and theoretical research has increased exponentially in every field of vision, from biophysics to the perception of color, form, and motion by humans. Some notions of Gestalt psychology and physiology once a pole apart have been joined in a mathematical union. Professor Corso has excluded this work from his book.

It may be said that what Professor Corso aimed for was an exposition of the experimental psychology of sensory behavior, not of sensory physiology. If that is true, he has misled us to believe that physiology should inform psychology: "this direct interdisciplinary attack on problems of behavior tends to minimize any inconsistencies which might otherwise arise through the study of psychology and physiology as isolated areas" (pp. 82-83). There is little mutual interplay between physiology and psychology in the text. A good deal of information is contained in the two chapters (3 and 4) on the nervous system and sensory receptors, but it is abstract and summary. Graphs and figures showing what is connected to whom could have been added to breathe life into these chapters. In any case, not much use is made of this neurology and physiology in later chapters, although the number of pages spent on abnormalities sets one to expect more. An illustrative case is the 40 pages on sensory deprivation (Chapter 14).

The physiological theory advanced to account for the phenomena of sensory deprivation amounts to saying that "the reticular system monitors all input and output activities of the sensory systems and, as a consequence, becomes adapted to certain levels of adaptation" (p. 588). Feelings and acts are influenced by the response level of the reticular system. When the normal level is exceeded, "the 'unusual' effects of sensory deprivation are obtained" (p. 588). No attempt is made to relate a single fact of sensory deprivation from the preceding 38 pages to this loose hypothesis.

Chapter 6, on experimental design and the testing of hypotheses, is, in my opinion, misleading in title and inadequate in content. It consists mainly of a section called "Some Representative Experimental Designs with a Single Independent Variable." What is actually covered? It discusses t tests for independent and correlated means, and randomized design with more than two groups; the factorial design is informally defined, but "we will not go into the details of this analysis" (p. 217).

I was struck by the great number of photographs (8 half pages and 2 full pages) of brain surgery: control rooms, instruments, transducers and implanted electrodes, and stereotaxic devices. Most of these give no hint about the technique or process. Simple diagrams could have been informative. The same applies to a photograph of a sound-level meter, and to one of an esthesiometer that works in mysterious ways while looking rather like a perceptron seizing a psychological moment. But against the many uninformative photographs there was no diagram showing the theory of spatial conductance of the nerve impulse, or of the digital-analog-digital process relating two axons to each other through a synapse. There were no illuminating electron micrographs showing, for example, the basic ciliary structure of rods, cones, and hair cells of the cochlea and labyrinth.

The best discussions occur in Part IV, where there are some very useful treatments of theories of sensory discrimination (Chapter 11), of information theory (Chapter 12) and of adaptation-level theory (Chapter 13). I have already suggested that "Sensory Deprivation" (Chapter 14) is misplaced in a section on theory. Except for the excessively long section on "the neural quantum theory of sensory discrimination," the discussions in all four chapters are quite uncritical. Some 47 pages are devoted to information theory, which is said to have "tremendous utility" (p. 506), an opinion that will be shared by few these days. I grant that the notion of uncertainty has been used in a masterly way by a few—most notably G. A. Miller and W. R. Garner—but it has serious limitations, as Professor Corso wisely cautions.

The Experimental Psychology of Sensory Behavior is in truth not a contemporary account. For one thing, it reflects the conservative attitudes of the 'hypothetico-deductive method' of 30 years ago, as may be seen in some of the "Orientation" sections with which each chapter begins. Professor Corso delicately shuns explanations of psychology in physiology, and physiology in physics (reductionism). Each science has its proper laws. Cross-fertilization may be too costly in time and effort (p. 31). "The unity of science which is sought in reductionism must be achieved through the use of a common methodology, not the anticipation of a common theory or set of laws" (p. 30). Add to this the notion of unidirectional dependency of fields "from left to right, with psychology placed between the physical-biological sciences and the social sciences. The dependency, however, is not shown to be among laterally arranged fields, but upon philosophy" (p. 26). After reading this it becomes easier to understand why Professor Corso did not often integrate neurology and psychology, and why he omitted, for example, linear-systems analysis of audition and vision; mathematical, computer, and cognitive (information-processing) models; mechanisms of sensory storage and retrieval, perceptual learning, and speech perception. Of course, some of these omissions may be related to the fact that only about 12% of the references are later than 1961.

It must be said that the book is attractive and well-made. The line drawings and graphs are excellent. It must also be said that the bibliography and author index contain many annoying errors caused by too-careless proofing. There is much useful, detailed material in Chapters 7 ("Psychophysical Methods"), 8 ("Absolute and Differential Thresholds"), and 9 ("Adaptation and Other Sensory Phenomena"); and in Part IV, "General Theories in Contemporary Experimental Psychology." By the careful addition of experimental and theoretical work, together with critical discussion, some instructors may be able to eke out the book into a useful text.

EDWARD CARTERETTE, *University of California, Los Angeles*

HELMHOLTZ ON PERCEPTION: ITS PHYSIOLOGY AND DEVELOPMENT. By RICHARD M. WARREN and ROSLYN P. WARREN. New York: Wiley, 1968. Pp. x, 277.

This book is a collection of six selections from Helmholtz's writings on perception. Publication of the book was undoubtedly stimulated by the recent renewal of interest in perception. However, only two of the selections, "The Facts of Perception" and the "Origin of the Correct Interpretation of our Sensory Impressions," have been translated by the editors for this volume. The other four selections have been available in English translation for many years. It is unfortunate that the editors did not choose to include translations from the second edition of the *Treatise on Physiological Optics*. The third edition, which was translated into English, is a reprint of the first edition with editorial addenda; but the second edition, which has never been translated into English, represents Helmholtz's own final thinking on many problems. The present collection is a useful source of Helmholtz's views on perception but contributes little that was not previously available to the psychologist who does not read German.

The book is, in a way, a hybrid. The first three selections are taken from

Helmholtz's *Popular Scientific Lectures*, published originally by Longmans Green in 1881 and reprinted by Dover in 1962. These explain, at a nontechnical level, Helmholtz's theories of pitch discrimination, consonance and dissonance, color vision, binocular perception, and brightness perception. The lectures are clearly written, with detailed descriptions, and are generally easy to follow. Helmholtz's work laid the foundation for much of the subsequent research in these areas. However, subsequent developments have led to major advances, which make these selections primarily of historical interest. The editorial notes are sometimes helpful in clarifying the relationship between the current state of knowledge and that of Helmholtz's time.

The second three selections deal explicitly with defining the concept of perception. They are technical, difficult, and directed to the specialist. In the selections "Concerning the Perceptions in General" and "The Origin of the Correct Interpretation of our Sensory Impressions," Helmholtz formulates his view of the perceptual process. In the selection "The Facts of Perception," Helmholtz's motivation appears to be philosophical rather than psychological. He is concerned with showing that the origins of our ideas are derived from experience, and he argues that if we can imagine non-Euclidian spaces, then Kant's contention that the axioms of Euclidian geometry are necessary consequences of a transcendental intuition is false.

Central to Helmholtz's theory of perception are two propositions. The first is that "we are not in the habit of observing our sensations accurately except as they are useful in enabling us to recognize external objects. On the contrary, we are wont to disregard all those parts of the sensations that are of no importance so far as external objects are concerned" (p. 172). The second central proposition is that "such objects are always imagined as being present in the field of vision as would have to be there in order to produce the same impression on the nervous mechanism, the eyes being used under normal conditions" (p. 175). Helmholtz accepted the distinction between the immediate, simple units of awareness (sensations) and the complex mental compounds that the mind builds from them (perceptions). Perceptions, however, are not an aggregate of sensations. For Helmholtz, sensations are signs or cues for external objects. The mind utilizes the information provided by sensations to infer the properties of external events and objects. By emphasizing that perception is directed toward objects rather than toward sensory qualities, Helmholtz lays the groundwork for the functionalist viewpoint. Perception is to be understood in terms of how it allows an individual to behave adequately with respect to his environment. Helmholtz's hypothesis that perception represents inferences formulated on the basis of sensory information has played a pivotal role in perceptual theorizing. Over the years, this hypothesis has been applied in many ways. The concept was enlarged by Brunswik in his theory of probabilistic functionalism, in which the perceptual system is conceived of as weighing fallible cues on the basis of past experience. In recent years the hypothesis has been given a sophisticated mathematical formulation in terms of statistical decision theory. If one substitutes for sensations the analytic operations that convert a stimulus into features or cues, and for judgments done by the mind computational like processes carried out by neural networks, Helmholtz's hypothesis underlies many current models of perception.

Though perceiving and thinking are similar in their hypothesislike character, there are substantial differences between the two. Perception, unlike thinking, involves processes that are rapid, automatic, and irresistible. It is the stereotyped nature of perception that is the basis for illusory perceptions such as occur in the Ames 'distorted room.' The room is seen as rectangular because this is the 'best bet' among alternative interpretations. Because the room is seen as rectangular, people are seen to grow and shrink as they move from one side to the other along the back wall. Once established, the frame of reference determines the perception of size even when the resulting percept runs counter to usual experience. Changes in the retinal sizes of an object become perceived as changes in the distances of the object only if the object becomes an invariant anchor and changes the spatial framework. An object's apparent size follows from the spatial framework established, and not solely from past experience with that object. The rapid, stereotyped nature of perception suggests that it depends on specific and automatic mediating mechanisms, and not on a general cognitive inference. This is consistent with numerous experiments showing that perception depends on highly specific information given visually.

Helmholtz was aware that there are important differences between perception and thinking. The inferential procedure in thought involves the formulation of rules, the examination of premises, the checking of inferences, and conscious intermediate stages. Perception involves intuitive inferences in which sensory data produce the perception of an object without clear and conscious intermediate stages. Though Helmholtz often speaks of a general inferential process without regard to specific mechanisms, he was also well aware that perceptions may depend on specific mechanisms. For example, he suggested that the stability of the environment during saccadic movement derives from a comparison of the perceptual feedback with the command signals producing these changes. There is a comparison of the displacement of the field with the initiating signals. If the signals resulting from retinal movement match the copy of the command, the environment is perceived as stationary. When the two do not match, motion is perceived. Knowledge that the eye and not the environment is moving, as when one jiggles his eye, still causes apparent movement. The information has to be given in terms of a centrally initiated eye movement and a specific visual feedback if the apparent movement is to be canceled. A second illustration may be taken from Helmholtz's treatment of color perception. Helmholtz's theory of color constancy is often described as involving a judgment in which the color of the illumination is discounted. This discounting, however, occurs in terms of a specific medium. The average color reflected from a scene defines an achromatic point. The discounting of the illumination is to be interpreted in terms of a shifting of the achromatic point in the color space. It is not knowledge of the illumination per se that is important; what is important is the specific information given to the visual system, which determines the achromatic point. Helmholtz's explanation of color contrast, often said to involve an error of judgment on the part of the observer, is also based on a shift of the achromatic point due to the color of the background.

We have made considerable progress in the study of sensory processes since Helmholtz. New experimental techniques as well as new concepts have been developed for the analysis of sensory functions. For example, quantitative measurements of retinal photopigments, the traveling-wave theory of pitch discrimination, and electrophysical and psychological studies of lateral summation and inhibition have provided a fundamental knowledge of the sensory mechanisms in audition and vision about which Helmholtz could only speculate. Unfortunately, we have not developed comparable new concepts and findings in the area of central processes. Although the development of computers has stimulated a number of theories of perceptual processing that have proven to be of interest for the problems involved in performing automatic shape and pattern recognition, these have not succeeded in elucidating the fundamental perceptual processes. To take a simple example, Helmholtz was aware that the continuous sequence of perspective projections of an object that accompanies head movements yields the appearance of a solid object from altered points of view. What are the rules by which this sequence of perspective views is mapped into a perceptual object? It is commonly ascribed to a general inference deriving from our experience in moving about space. But an explanation in terms of past experience does not explain the precise coupling between movement of the object and movement of the head that occurs in the Mach-Eden illusion. In the Mach-Eden illusion a card is folded into the form of a W and placed on a table. When it is viewed with one eye, one is able to see the card in false perspective standing upright on the table. When it is seen as upright, the perspective projections of the card accompanying movements of the head cause the card to appear to twist and turn. This illusory movement shows the power of Helmholtz's rule that we see objects as they would have to be under normal conditions in order to produce the prevailing retinal image. However, the mechanism establishing a precise coupling between the transformations of a retinal image and the movements of the card is unknown. The challenge to perceptual theory is to discover the mechanisms that explain the basis for Helmholtz's empirical rule for predicting perception. We have few hypotheses about the mediating mechanisms underlying the perception of a veridical, stable, and structered world that go beyond Helmholtz's classic work.

Jacob Beck, *University of Oregon*

INFORMATION THEORY AND ESTHETIC PERCEPTION. By Abraham Moles. Translated by Joel E. Cohen. Urbana: University of Illinois Press, 1966 (cloth), 1968 (paperback). Pp. xi, 221. $7.50; $2.45.

This book, first published in France in 1958 as *Théorie de l'information et perception esthétique,* is part of the enthusiasm for information theory that peaked in the 1950s. Professor Moles connects the ideas of information and communication theory to problems of aesthetics (especially to those of music) in an imaginative but very general way. After more than a decade, much of it filled with reaction to and criticism of such extensions of information theory, it is inevitable that many parts of the book seem naive and dated. Nevertheless, in spite of its weaknesses, this book is an important attempt to bring some kind of scientific theory and empirical base to the problems of aesthetics. Its emphasis on the analysis of the structure of the aesthetic stimulus is especially good. Anyone familiar with the existing mishmash of subjective inconsistencies loosely called aesthetic theory knows how much we need any approach that brings even the smallest amount of rigorous thought and empirical leavening.

Now that the book is out in paperback, it should be a useful addition to any contemporary course that approaches aesthetics from a scientific viewpoint or studies those contemporary art forms strongly influenced by science and technology (electronic music, computer art, and so on).

PAUL VITZ, *New York University*

LEARNING: PROCESSES. Edited by MELVIN H. MARX. New York: Macmillan, 1969. Pp. 573. $9.95.

This is the first of a series of three books intended as texts for an upper-level course in learning. The other two books, *Learning: Theories* and *Learning: Interactions,* are to appear later. The present volume is divided into sections, each section by a different author and each containing one or more chapters. This approach has the advantage of allowing each author to write solely in the area of his expertise. Most of the sections are comprehensive up-to-date summaries of their fields. The body of the work—sections by D'Amato on instrumental conditioning, by Gormezano and Moore on classical conditioning, by Kalish on stimulus generalization, by Cofer on verbal learning, by Ellis on transfer and retention, and by Adams on motor behavior—accomplish what the editor has intended. The book covers what a college senior graduating with a major in psychology should know about the subject.

The approach, however, has disadvantages. There is some variance in style. For instance, Gormezano and Moore's chapters contain much more detail and are more empirically oriented than D'Amato's and Kalish's chapters, which, despite the empirical objectives of this volume, spend many pages on theoretical questions. Given the nature of the material, such differences in style may be unavoidable, but a single author would have minimized them. A more serious deficiency is that the approach allows the existence of gaps. While each area covered by the several authors is covered well, there are some areas that are only briefly touched upon. There is an entire section (Kalish's) on gneralization, but there is very little on secondary or conditioned reinforcement. In fact, this topic is not mentioned in the index. Choice, probability learning, and attention receive only cursory treatment. A short chapter (Marx's) on the free operant discusses shaping and schedules of reinforcement; but it is far below the other sections in terms of comprehensiveness and quality.

The excellent detailed treatment of several specific areas makes this book suitable as an advanced-level text, but because of its several gaps, it could not serve this function alone. Most instructors would want to use much supplementary material with it or to use this volume as a supplement to another text.

HOWARD RACHLIN, *State University of New York at Stony Brook*

THE MOLECULAR BASIS OF MEMORY. By EDWARD M. GUROWITZ. Englewood Cliffs, N. J.: Prentice-Hall, 1969. Pp. viii, 89. $3.50.

This book is essentially a critical review and discussion of the complex area relating macromolecules to memory and learning. It is clear and complete (up to January, 1968). No particular background is needed; a short chapter on the biochemistry of RNA and protein is included for the reader unfamiliar with these concepts of molecular biology. The author begins with a discussion of the hypothermia and electroshock experiments that form the basis for the belief that memory is not dependent entirely on ongoing electrical events but has, in addition, a structural or chemical basis. The next two sections describe and analyze the evidence that RNA and protein are the molecular substrates of memory. The author concludes that both molecules are probably involved but that the hypothesis does not rest on any really firm experimental grounds. Rather, the attractiveness of the hypothesis comes from a battery of suggestive experiments. A few sections on molecular theories of memory are contained within these chapters. These sections, in my opinion, are neither comprehensive nor particularly penetrating. The final section describes the controversial interanimal transfer experiments. It is interesting to read through the maze of experiments and hypotheses, but it is really impossible to draw any conclusions. The author's conclusion that "data seem to indicate that the inter-animal transfer phenomeon is a real, albeit highly elusive, effect probably based on a polypeptide or small protein" (p. 76) is, it seems to me, premature.

I rather enjoyed this book and found it readable and quite complete. I would recommend it as an introduction to the field or as a refresher. I did not find it particular strong in presenting an overview of the nature of the problem, however, or in highlighting certain key experiments or their conclusions. Because the field is developing so rapidly, the interested reader will need to supplement this book with recent studies from the literature. This is particularly true for studies of memory using protein-synthesis inhibitors and of RNA changes found with training.

Carl Cotman, *University of California, Irvine*

READINGS IN PHYSIOLOGICAL PSYCHOLOGY: NEUROPHYSIOLOGY/SENSORY PROCESSES. Edited by Charles G. Gross and H. Philip Zeigler. New York: Harper & Row, 1969. Pp. viii, 286. $5.25.

This paperback is one in Gross and Ziegler's series of three collections of readings in physiological psychology. The other two collections cover the areas of motivation and of learning and memory. The books are designed for undergraduates, and can be used either separately or in combination. The readings in neurophysiology and sensory processes are meant to illustrate "the relationship between basic neurophysiology, sensory physiology, and behavior" (p. 1), and consist of 15 reprints of articles from usually accessible professional journals. Two of these concern neurophysiology, and the remaining articles cover sensory processes, grouped according to the following topics: receptor mechanisms, sensory coding, and central control of sensory input.

Given the rate at which collections of readings are being published, it seems appropriate to make a few evaluative comments concerning this practice. First, no book of readings is entirely satisfactory to any instructor. Because of differences in emphasis and differences of opinion about the importance of specific articles, individual instructors prefer different sets of readings. Second, undergraduates often find journal articles difficult, and indeed, discouraging, for a variety of reasons. Some students are overwhelmed by detailed or lengthy methods sections, and some by esoteric discussion sections. In fact, many students entirely miss the significance of journal articles. This is related to a third possible criticism of books of readings. Editors in general supply little or no introductory material that would prepare the student to read an article and help him to realize its contribution to the field.

It is certainly not my contention that undergraduates cannot or should not read journal articles. On the contrary, I think journal articles can be very worthwhile for undergraduates, provided that these students are prepared to read them. One such preparation can be provided by class lectures. Another could be the preparation provided by editors, in the form of explanatory remarks before and after each article. Gross and Zeigler supply this to some extent. Before each group of readings are introductory remarks that clearly place the papers in historical perspective and indicate the overall goal of the research. This material is very well presented and more extensive than that usually found in collections of readings. But even here it is not sufficiently detailed, and in fact comprises merely 6% of the total text. Finally, in some rare cases, articles need no student preparation for them. That is, some articles provide a more general treatment of a large body of research and clearly state the procedure and goals of this research without esoteric detail. Some invited addresses to the American Psychological Association fall into this category. Although a few of the readings selected by Gross and Zeigler are of this type, the majority are not.

Thus, to conclude, I would not select this book for my course, but this decision is based on the above criticisms of books of readings in general. However, many psychologists do not share a number of my opinions. In fact, to be fair to these editors, it is probably true that many instructors of sensory psychology and sensory physiology will be satisfied with most of their choices and will find their explanatory material more extensive and more informative than is usually the case in such collections, and consequently will want to consider this text for their courses.

Donald C. Hood, *Columbia University*

SPEECH BEHAVIOR AND HUMAN INTERACTION. By George A. Borden, Richard B. Gregg, and Theodore G. Grove. Englewood Cliffs, N. J.: Prentice-Hall, 1969. Pp. viii, 260. $5.95.

Drawing concepts and experimental results from a variety of sources including neurology, psychology, and cybernetics, the authors attempt to focus a wide range of knowledge in terms of speech behavior in human interaction. The book is intended as a text for introductory courses in speech, though certainly not for traditional courses concerned with how to make effective speeches. The first part of the book deals with processes of communication within the individual; the second part examines interpersonal communication; and the final section deals with public communication. Unfortunately, the three major sections are not very well integrated, each representing a more or less independent essay with only superficial cross-references from one section to another.

Having taken my first and last speech course some 25 years ago, I was delighted to discover that this text was designed to offer students something other than recipes for making after-dinner speeches. But while the authors' aims are entirely laudable, they have achieved only limited success in their present text. As a result of trying to cover an extraordinary range of knowledge within the limits of a relatively short book, they are forced to treat most topics in rather general, and often vague, terms. Thus, the serious beginning student may well reject this approach to the study of speech behavior, feeling that the contributions of other disciplines are too superficial to warrant further intensive study. Moreover, the authors sometimes become involved in topics only tangentially related to speech behavior, and the reader primarily interested in speech may easily lose interest as the text moves from a consideration of self-concept, to role theory and a discussion of public communication that goes far beyond the usual definition of speech behavior.

In general, then, this text represents a desirable movement in the direction of integrating relevant disciplines in the study of speech. But in order to achieve this integration meaningfully, a much more systematic and intensive treatment is required than that offered by this book. Perhaps the authors might profitably view this volume as an outline for subsequent work that will do justice to their admirable aims.

JOEL R. DAVITZ, *Teachers College, Columbia University*

HUMAN INTELLIGENCE: ITS NATURE AND ASSESSMENT. By H. J. BUTCHER. New York: Barnes & Noble, 1968. Pp. 343. $8.75.

This is an important book. Psychology's increasing preoccupation over the last three decades with learning theory and experimental design has tended to obscure the empirical and descriptive areas known collectively in an earlier day as 'individual differences,' within which was 'intelligence.' Yet 'intelligence' and its measurement remains psychology's firmest accomplishment. No other psychological dimension has been more effectively measured, from the standpoint of stability and prediction to performance, and certainly no other contribution has had comparable immediate and long-range effects on society. Yet 'intelligence' is generally neglected, even dismissed as meaningless, in contemporary psychology.

The author states his explicit aims: (*a*) "to treat the subject of human intelligence not as a narrow specialty but as a fundamental topic in psychology, to which many different kinds of survey, experiment and speculation are relevant"; (*b*) "to provide an up-to-date guide to much of what is known, and where it can be found with particular reference to work done in the last decade"; (*c*) "to swing the pendulum a little against the current state of opinion, insofar as this holds that 'intelligence' is an out-dated concept" (p. 9). These objectives are brilliantly achieved. At a time when scholarship in psychology consists of reports of programmatic research, much of it in the idiosyncratic terminology of particular investigators, sometimes overextended and often overstated in theoretical language, a review of the literature that dispassionately and judiciously sets in order a particular concept, evaluates it, and shows what particularly needs to be done, is especially valuable.

The author's argument makes clear the atheoretical nature of the concept of intelligence and what must be done to give it a better theoretical position. He is committed to a positivistic position, yet flexibly so. He recognizes the fallacies of hypostatization into which the noun 'intelligence' has led us and proposes that we consider the adjectival or adverbial forms of the word as more useful. He proposes, and defends with considerable success, the concept of a general ability as set off against more qualitative ways of viewing functioning, marshaling the evidence that we here deal with a dimension that has both statistical-logical and functional-psychological meaning. He recognizes the dangers inherent in 'open' concepts not fully specified operationally and argues persuasively for progress toward the latter goal. But he does not want to commit us prematurely to arbitrarily selected and limited operations. His discussion of the definition of intelligence deserves to be read both by those who would push us precipitously into operational definitions and by those who are reluctant to specify too sharply less some qualitative 'richness' of meaning be lost. He offers an excellent statement of some epistemological issues in psychology's 'philosophy of science.' In treating conceptual processes, the author makes a practical distinction among several terms now used in rather various ways (p. 78). 'Concept learning' is equated to the mental processes employed in experiments on discrimination learning; 'concept attainment' and 'concept identification' are synonymous and describe the results of experiments in which the subject is required to identify a concept predetermined by the experimenter. The term 'concept formation' he reserves for the kind of work more appropriately done with children's thinking, especially by Piaget and his colleagues. The author recognizes the complexity of these issues, their resistance to analysis into simple variables capable of manipulation in the laboratory, the interactive aspects of many components of the concept or construct, and the extreme difficulty of setting up designs that will appropriately allow for these complex interactions. He is optimistic about ultimate solution of these problems but believes that this may take much time and great patience. He emphasizes that the topic under consideration is really more than a study of behavior; it is a study of one of the most characteristic and distinctive properties of man's behavior—intelligent behaving.

A third contribution of considerable value is an extremely lucid, though admittedly brief, introduction to the simulation of intellectual behaviors by computers. This is no more than an introduction, but it is extremely well written and comprehensive. The author refers to a large number of substantive and research papers, and outlines clearly the area in which we may expect substantial contributions in the next decade.

A fourth substantive contribution is found in two chapters relating to the development of normal intelligence and to the influence of heredity and some related problems. His presentation of the classic argument, its development over the years, and its present state is succinct. An extremely important contribution is his recognition of the social and political significance of this problem. With astonishing impartiality he clarifies the extent to which investigators' political and value commitments have influenced their studies. On the heritability of intelligence he reaches several conclusions and evaluates accurately the rather shaky character of some of the evidence. First is that differential language experience in the early years certainly cannot account for all the variability that is found. Another is that a completely culture-free intelligence test is impossible. A third is that the deleterious effects of unfavorable environments may be greater than the beneficial effects of good ones, especially on children with above-average levels of ability. Such a conclusion seems reasonable but is supported only by indirect evidence.

The discussion of the role of personality in achievement is extremely useful. Prediction equations are used in virtually every major university and college in the United States and widely used in England. The author makes very clear that as prediction succeeds and improves selection, the predictive indexes must necessarily fall. Within groups highly selected according to abilities, the contribution of personality factors to performance and achievement becomes proportionately greater. He urges more attention to these and includes the important work presently being done on these variables.

A final contribution of particular value to American readers is the discussion of the 'streaming' approach in British education at age 11. The author sees this practice as a distinct step forward over the selection to higher education that had existed prior to the mid-1940s, but he also sees that this solution is far from ideal. It has not broken through the class structure, permitting free access of capable individuals from lower social classes to higher education, as had been hoped. His discussion recognizes that with all the criticisms leveled at the concept of general education in the United States, there is much to be said for that type of solution in a truly democratic society. His discussions of the significance of prediction and selection by tests have considerable relevance to us in the United States at this time. They are wise and judicious, and yet clearly reflect a commitment to the ideal of generous educational opportunity for all capable individuals.

DALE B. HARRIS, *Pennsylvania State University*

SOCIAL PSYCHOLOGY AND HUMAN VALUES. By M. BREWSTER SMITH. Chicago: Aldine, 1969. Pp. x, 438. $12.50.

Brewster Smith has collected in this volume 28 of his own essays. Of these 28, 4 have not been previously published, 9 have appeared in different books, and 15 have been spread over 11 different journals. The impression that the content of these papers must reflect a heterogeneity befitting their origins is reinforced by an initial inspection of their titles. Proust, Peace Corps teachers in Ghana, family planning, and Samuel Stouffer, to say the least, are unlikely bedfellows. Surely in this age of methodological complexity and ever-increasing specialization, no one could seriously publish a book whose subject matter crossed the boundaries between psychology, anthropology, and political science, and at the same time present a bold new approach to mental health care, transmitting new articulations of the standards governing professional and scientific responsibility. But Smith does. When, in addition, the same author presents original data on authoritarianism, the morality of student activists, and racial prejudice in schools, together with an analysis of the concept of competence that makes White's conceptualization look almost short-sighted, then one must pause. Perhaps the author might, in a not-so-obvious way, be qualified to write encyclopedic comments on the self and cognitive consistency as well as on attitude change. Brewster Smith does all this and more. And he does it well.

Persevering readers will be intrigued by the sharpening conceptual focus on personal values, political rationality, positive mental health, competence, the reflexive self, and freedom as legitimate (and rather closely related) objects of empirical, psychological inquiry. From such a perspective one might begin to share the view, expressed in Smith's title essay, that other social psychologists these days are somewhat preoccupied with mere experimental ingenuity.

HENRY ALKER, *Cornell University*

THE STUDY OF POLICY FORMATION. Edited by RAYMOND A. BAUER and KENNETH J. GERGEN. New York: The Free Press, 1968. Pp. xxii, 392. $9.95.

If you read these nine chapters in reverse order of their presentation, you are likely to conclude that this is a more important book than if you read it in the conventional front-to-back direction. This is so because the final section contains three of the best written, most insightful case studies on the complexities of policy making that are available. They demonstrate with clarity what the earlier chapters seem to obscure, namely, the precise definition of the domain of inquiry that is encompassed by the phrase "the study of policy formation."

The purpose of the book is to provide a "series of position papers that would furnish background for the design of empirical research on the process of policy formation." The six chapters that are not case studies consist of an attempted 'unification' chapter (Bauer), two extremely thorough and readable reviews of normative economic theory and descriptive decision theory (the former by Zeckhauser and Shaefer, the latter by Bower), a discouraging and largely redundant 'state of the art' chapter (Schoettle), and two chapters that represent a thoughtful effort to turn the sprawl of policy making into researchable components (Gergen).

There is much here of interest to psychologists, although one must wade through considerable repetition and ponderous prose to find it. There is a glaring discrepancy between the abundance of concepts to explain individual choice behavior and the scarcity of those to handle collective choice. The seriousness of this latter deficit is made all the more apparent by the frequency with which occasions of collective choice occur in the case studies. Those who study groups and intergroup relations clearly have more to offer the topic of policy making than they may realize. The variable, and reality, of self-interest is discussed with refreshing candor. This helps to counterbalance the sometimes hyperaltruistic assumptions about group action that psychologists are prone to make. The concept of "leverage points," a conceptual and empirical wedge into policy making proposed by Gergen and Bauer, has the dual attractions of being consistent with what is known about interpersonal influence processes and of being workable.

More must be said about the case studies, since they alone warrant purchase of the book. They concern the topics of urban mass transportation, technology transfer (adapting defense R and D to civilian uses), and foreign aid. Listing the topics under these mundane labels is actually misleading, since the amount of life and light the authors inject into these somber topics is awesome. Each study has its own distinctive assets. The urban-transportation study (Schneider) benefits because the author works within a model that fits rather than obscures what happens. An advantage of Furash's technology-transfer chapter is that it provides the scientist who knows little about policy making a set of variables and issues with which he is already familiar. This chapter also has an unusual amount of contemporary relevance, given the successes in space exploration against a background of questions concerning earthly payoffs, and given the mounting concern about the 'relevance' for society of what academics do in the Ivory Tower. Geiger and Hansen's discussion of decision making in foreign policy is perhaps the most stunning of the three studies, simply because of the way the authors are able to keep under control and interweave the incredibly diverse and complicated interests that converge on this form of policy making. I can think of no more sobering or challenging way to confront students with the realities of doing social science than to immerse them in this case and then ask, 'OK, now how do you study *that*?' And that is one reason why I think this very long book is in fact too short. What it needs is an appendix written by the authors of the case studies telling us how *they* did it. To uncover and organize as much as these authors did, they must have used resources, intuitions, 'techniques,' that eluded the writers of the other six chapers. They must have done what, one would conclude after reading the earlier material, cannot, or has not, been done.

The case studies contain embarrassments, not only for theorists and methodologists at large, but even for the contributors to this book. For example, Gergen, developing the concept of 'issue relevance,' states that "an issue will be relevant to an individual to the extent that for him it can potentially modify the status quo. For example, changes in urban mass transportation would be of less relevance to the farmer than to the urban department-store owner or commuter" (p. 183). And yet, when we turn to Schneider's discussion of the hearings that were held prior to the enactment of the Urban Mass Transportation Act of 1964, we discover (p. 264) that the American Farm Bureau Federation gave strong testimony against the enactment of the legislation with the argument that "urban mass-transportation aid would add to the unhealthy trend for federal expenditures to increase relative to Gross National Product." Even more telling, the Farm Bureau's position was identical with that of the United States Chamber of Commerce, among whose members would surely be found Gergen's "urban department-store owner." The case studies also offer substantial support for a formulation that is given little attention in the volume: Braybrooke and Lindblom's "strategy of disjointed incrementalism." The formulation is summarized on one page, and the restatement loses many of the crucial nuances that were present in the original. One would expect that Bauer would pick up themes such as this in his overview chapter, since it was written after he had all the papers in hand. In fact, he does not, and this is one of the disappointments. It is clear from the way Bauer handles his chapter that he has more to say about policy making than he says. The regret is that there is more in those papers than meets his eye.

And therein lies the value of this book. It furnishes the resources from which one may be able to construct a more robust picture of policy-making processes than any one contributor was able to do. And since some of the contributions tempt one to do this, then it must be concluded that as a set of 'position papers' this collection is largely successful.

KARL WEICK, *University of Minnesota*

DISEASES OF ATTENTION AND PERCEPTION

Monte Jay Meldman, M.D.
Director of Research
Forest Hospital
Des Plaines, Illinois

241 pages integrating data on attention and perception in behavioral pathology into a classification system of diseases of attention. 580 references review the important literature.

Available from your local book store and from
Pergamon Press, Inc.
Maxwell House
Fairview Park
Elmsford, New York 10523

Price: $13.50

BINOCULAR VISION: A PHYSICAL AND A NEURAL THEORY

George Sperling, *Bell Telephone Laboratories, Murray Hill, New Jersey* and *New York University*

Abstract. The classical model of energy levels consists of a bumpy surface upon which a marble rolls freely. Dips in the surface represent stable states in which the marble can become entrapped. This physical theory is applied to three major processes of binocular vision: accommodation (a), horizontal vergence (v), and fusion (u). A simple model illustrates how the same external stimulus may produce different vergence states, depending on the preceding stimuli, and it shows how extreme values of vergence can be achieved. The model applies directly to vertical vergence and to torsional eye rotations. A similar model fits the accommodation and fusion systems and can be extended to account for the v-a-u interactions.

The neural structures underlying the systems differ. The neural theory of fusion proposes two neural binocular fields (NBFs): a primary NBF for fine details and for fine depth discrimination in the stimulus (its outflow corresponds to the cyclopean binocular view); and a secondary NBF for coarse details and large depth signs (its outflow directs fine analysis to relevant portions of the visual field, it contributes nothing except coarse depth signs directly to the cyclopean view). The fusion model illustrates how the same retinal stimulus may produce different stable perceptual states, depending on recent stimuli, and it accounts for fusion and rivalry within the same system and the same neurons. Appendix A offers a general definition of image blur, and Appendix B a quantitative analysis of the multistable phenomena in rivalry and fusion.

Because the two eyes provide two different views of the world, a person with normal binocular vision can determine the relative depth of unfamiliar objects in his field of view when he and the objects are motionless. This ability is called stereoscopic depth perception—or, often, just binocular depth perception, to emphasize its dependence on both eyes. The evolutionary value of binocular depth perception is obvious, and some interesting and complicated adapta-

---

Received for publication May 5, 1970. The author expresses his appreciation to Dr. Mohon Sondhi for his help with Appendixes A and B, to Dr. John Krauskopf for his stimulating criticisms, and to Dr. C. F. Strohmeyer III and the late Dr. Lloyd Marlowe for their helpful comments on the manuscript.

Date... 9/5/71

tions have evolved in our visual systems to facilitate it. This article explores the intricate interrelationships between the three most important mechanisms for binocular vision: the mechanisms for vergence, for accommodation, and for fusion.

The most important subjective phenomenon of binocular vision is binocular *fusion.* When corresponding points of the retinas of the two eyes are stimulated with light from the same (or nearly the same) object, only one object is seen. We say the two retinal images are fused. Ordinarily, the experience of one object is so compelling that fusion requires no further definition.[1]

Binocular fusion is possible only when stimuli fall on corresponding—or nearly corresponding—points of the retinas. To clarify what we mean by corresponding points, we suppose the eyes are looking at a point at infinity. (For practical purposes, this means a point on an object several hundred yards or more away.) Then, the lines of sight of the left and of the right eye are parallel, and identical images fall on each retina. For the moment, we define corresponding points as the corresponding retinal points under these retinal images.[2] To view a near object the eyes must rotate inward (converge) in order for the object to stimulate these corresponding points. The *vergence* movements of the eyes are the second remarkable adaptation of binocular vision. When the eyes are verged on a particular object, only objects in a limited range of depth planes can be fused; objects outside this range are unfused.

---

[1] For the purist, binocular fusion can be defined objectively in terms of shape (or pattern) invariance. The principle is that small lateral translations (in opposite directions) of the stimuli to the two eyes leave the shape of a fused pattern unchanged. Additionally, it also must be shown that both eyes' stimuli contribute to the perceived pattern. These are sufficient (but not necessary) conditions. Specifically, a pattern of illumination to the left retina $[l(x, y)]$ and to the right retina $[l'(x, y)]$ are binocularly fused if there exists a Δx, $|\Delta x| > 2$ min., such that when the new pattern $[l(x + \Delta x, y)$ and $l'(x - \Delta x, y)]$ replaces the original pattern on the retinas, the observer reports the shape of the new pattern is the same as that of the old. Its depth in space may or may not have changed. To ensure that both l and l' contribute to the shape of the pattern, there must be points on the left stimulus [e.g., bright spots of light $\delta(x_1, y_1)$, $\delta(x_2, y_2)$ . . .] and on the right stimulus $[\delta'(x_1', y_1')$, $\delta'(x_2', y_2')$ . . .] such that the two sets of points $[\delta(x_i, y_i)$ and $\delta(x_i', y_i')]$ are all different and the observer reports seeing points from both sets in any neighborhood where fusion is being tested. The motivation for this definition derives from the neural model of fusion.

[2] This actually is a definition of 'corresponding coordinates.' Subsequently, 'corresponding points' are defined as those whose projections intersect in the middle layer of the neural binocular field. In the neighborhood of fixation, corresponding coordinates and corresponding points coincide exactly.

The third major phenomenon of binocular vision is *accommodation,* whereby the lens of the eye varies its focusing power so that the depth plane of the object being viewed is in focus on the retina. Though accommodation is also characteristic of monocular vision, it is important in binocular vision because it is closely linked to vergence. That is, accommodating for a near object causes convergence, and convergence causes accommodation.

A priori, neither fusion nor vergence nor accommodation would be necessary for binocular depth perception. However, once a mechanism of fusion has developed, the evolution of accurate vergence and accommodation would enable the fusion mechanism to evolve further so as to substantially increase its power of depth discrimination. In fact, once all three mechanisms exist, there are such great advantages to binocular vision to be gained from their mutual interactions that these interactions are virtually predetermined. This paper presents two kinds of models, physical and neural, to account for the fundamental phenomena of vergence, accommodation, and fusion, and for the interactions between the three mechanisms.

Multiple stable states and path dependence. Our starting point is an easily observed phenomenon: the existence of *multiple stable states.* For example, Helmholtz, writing in the last century, observed that "the eyes also may be made to diverge by viewing stereoscopic pictures and gradually separating them farther and farther apart, all the time trying to fuse them into a single image. I am able in this way to produce a divergence of the lines of fixation of my eyes amounting to as much as eight degrees."[3]

To restate Helmholtz's procedure in more formal terms, he presented a picture (L) to the left eye and an identical copy (L') to the right eye by means of an optical device, such as a Wheatstone stereoscope. By moving L slowly to the left and L' to the right, he was able to cause his eyes to follow the movement of L and L', and thereby to diverge. But if L and L' had first been presented at the extreme separation of 8 deg., his eyes could not have diverged upon them; his eyes would have remained approximately in their initial position. This demonstrates multiple stable states of vergence. When L and L' are 'separated' by 8 deg.—that is, when the vergence angle required for fusion is 8 deg.—then the eyes may sometimes verge correctly and at other times be in their neutral position. There are two stable states of vergence for the same external stimulus. The particular vergence state attained by the eyes depends not only on the present stimulus but also on the sequence of previous stimuli.

---

[3] H. von Helmholtz, *Treatise on Physiological Optics,* 3d ed., J. P. C. Southall (trans.), 3, 1924 (reprinted 1962), 58.

As we shall see, Helmholtz's demonstration involves more complications than are apparent at first. First, the amount of divergence that can be attained depends on the choice of stimulus pictures (*L* and *L′*); for example, a greater amount of divergence is possible when pictures contain many high-contrast details. Second, with practice, some voluntary divergence is possible. Whether or not voluntary divergence suffices to produce fusion depends on the amount of divergence that is required and on the particular pictures (*L* and *L′*) being viewed. Third, fusion of *L* and *L′* can occur even when the eyes are not fully verged on *L* and *L′*. The first two complications, and of course the basic observation, are shown below to be immediate consequences of the physical model of vergence. The third complication follows from the joint action of the models of vergence and fusion.

It often is convenient to represent stimuli or responses as points in an *n*-dimensional space. In the example above, the vergence response can be represented in one-dimensional space as a point along a line, where the distance of the point from the origin represents the vergence angle of the eyes. The sequence of responses (to a sequence of stimuli) can be thought of as a path from point to point in this space. The conventional way of saying that a response depends not only on the present stimulus but also on the sequence of past stimuli is to say that the response is *path dependent*. That is, how the eyes respond to a particular stimulus (whether they verge or not) depends on their current vergence state as well as on the current stimulus. Because the current vergence state depends on previous stimuli, the response depends on both current and past stimuli.

In all the examples in this paper, path dependence and multiple stable states refer to the same phenomenon. The only way to achieve multiple stable states (different responses to same stimulus) is by taking different stimulus and response paths prior to presentation of the same final stimulus; the only way to demonstrate path dependence is to achieve different stable states to the same final stimulus (by virtue of having taken different paths).

Thus, multiple stable states and path dependence are logically equivalent.

Multiple stable states also occur in the fusion process. For example, suppose that vergence of the eyes is artificially maintained on a particular depth plane by optically canceling vergence motions.[4] Fusion is possible not only of objects in that particular depth plane but also of objects in other depth planes. The fusion or nonfusion of a particular object in a particular plane exhibits the same kind of multiple stable states as the vergence motions of the eye. That is, whether the object is seen as fused or double depends on whether it was seen as fused in its previous position, and so on. Finally, by methods to be described later, we may also observe multiple stable states of accommodation (focusing of the eye).

The multiple stable states of vergence, fusion, and accommodation follow from the same elementary principles, are representable by the same kind of physical model, and exhibit analogous correlated phenomena. Although the physical models of vergence, fusion, and accommodation are formally the same, the underlying mechanisms are different. The last main section of this article contains a speculative suggestion toward a neural model of these mechanisms.

THE PHYSICAL MODEL—WITHOUT INTERACTIONS

The classical model of energy levels consists of a bumpy surface, the *energy surface*, upon which a marble rolls freely. Dips in the surface are called *energy wells*; they represent stable states in which the marble can be trapped. Here, this classical model is applied to the three major processes of binocular vision: vergence (v), accommodation (a), and fusion (u). In these applications, the lateral position of the marble represents the instantaneous value of the depth plane of v or a or u. For example, the basic surface governing horizontal vergence is bowl-shaped, the shape being determined by internal factors. The marble rolls toward the center of the bowl, representing the tendency of the eyes to verge to a neutral position in the absence of a stimulus to vergence. The external stimulus is assumed to add perturbations to the basic surface, thereby creating new stable states.

---

[4] D. Fender and B. Julesz, Extension of Panum's fusional area in binocularly stabilized vision, *J. opt. Soc. Amer.*, 57, 1967, 819–830.

The physical model of binocular vision is thus based on a concept that, because of the physical analogy, is called energy (e). For example, in the case of vergence, the visual system is assumed to compute the net vergence energy of various possible states of vergence and—within certain limits—to achieve the state which has the lowest net vergence energy (e_v).

Let v denote the vergence angle between the eyes, a their accommodation, and u the plane of fusion. Let l and l' denote the stimuli on the left and right retinas respectively. There are assumed to be two component sources of energy, which add to produce net energy: *displacement energy* (g) and *image-disparity energy* (h). Displacement energy depends only on the current states of v, a, and u. It represents the amount of energy needed to displace v, a, and u from their natural resting values to their current values. Image-disparity energy depends only on l and l'. It represents the energy that a visual stimulus can contribute to displacing v, a, and u.

The physical model is a particular embodiment of a set of equations that expresses the general principles outlined above. The equations deal with the quantities e, g, and h, which represent respectively the *total*, the *internal*, and the *external* components of energy. Here the subscripts v, a, and u indicate the system whose energy is under consideration; and the variables v, a, and u represent the current values of vergence, accommodation, and fusion. Commensurate units for v, a, and u themselves are defined in the following sections. The equations contain all the interactions of the model and are presented here to provide an overview. They are not necessary for understanding the principles of the model, however, and readers who are not interested in mathematics should simply skip to the next section, on the model of vergence.

$$e_v(v, \cdots ,) = g_v(v - a) + h_v(l, l') \qquad [1]$$

$$e_a(a, \cdots ,) = g_a(a - v) + h_a(l, l') \qquad [2]$$

$$e_u(u, z, \cdots ,) = g_u(u - v, z) + h_u(l, l', z), \qquad z \subset R \qquad [3]$$

Equation 1 represents vergence; Equation 2 represents accommodation; and Equation 3 represents a set of equations for fusion, one for each region (z) of the retina (R). Each equation may be thought of as describing an energy surface upon which rolls a marble, whose projection on the v, a, and u axes respectively represents the value of v, a, and u.

Equations 1–3 do not give information about the dynamics of v, a, u, because time does not enter into them. Dynamic specification of v, a, and u requires another set of equations. Let x_1, x_2, x_3 represent v, a, and u respectively. The dynamics of x_i are defined by the first-order differential equations:

$$k_{1i}\frac{d^2x_i}{dt^2} + k_{2i}\frac{dx_i}{dt} = \frac{-\partial e(x_i)}{\partial x_i} \qquad (i = 1, 2, 3). \qquad [4]$$

The right-hand term of each Equation in 4 represents force acting on the marble; the coefficients k_{1i} and k_{2i} respectively reflect the marble's mass and the friction of the medium.

Although only the spatial derivative of energy (force) enters into Equations 4, the ensuing description of the systems is in terms of energy because it gives better intuitive insight. We may now begin to deal with the explicit definition of the quantities in Equations 1–3, the interpretation of the resulting equations, and their comparison with data.

Model of Vergence

Units of vergence. The horizontal vergence of the eyes is best expressed as the *angle* between their lines of sight. The angular unit of vergence is particularly convenient because it fits in with the spherical coordinate system used to describe the position of stimulus points in space and with the dioptric system used to measure accommodation. However, binocular vision is concerned mainly with depth, so it is useful to keep in mind that there is a one-to-one correspondence between the angle of vergence and the distance from the eyes to the verged-upon point. (We confine ourselves here to points directly in front of the observer, so-called symmetrical vergence.) The angular measure of vergence can thus be considered to be a transformed measure of depth, and in the discussion that follows, the term vergence *position* is used to evoke the idea of the distance at which the eyes are verged. The sign convention followed is that *con*vergence angles (lines of sight intersect in front of the observer) are positive and correspond to positive distances; *di*vergence angles (lines of sight intersect behind the observer) are negative and correspond to negative distances (i.e., behind the observer). In the case of divergence, of course, 'distance' does not have the usual physical interpretation, and we have no need of the distance transformation.

In this section on vergence, we assume that accommodation of the eye's lens is fixed, at a value corresponding to accurate focus of objects at the given viewing distance (e.g., 1 m.). To a first approximation, accommodation does remain fixed if stimuli are confined to a particular viewing distance (e.g., 1 m.). To maintain accommodation more accurately at a given nominal value would require a negative-feedback optical device.[5]

Vergence-displacement energy (g_v). For any given amount of accommodation of the eye lens, there is a natural vergence angle (v_0) of the eyes. That is, the natural vergence angle is the angle formed by the lines of sight of the eyes in the absence of any stimulus to vergence. (For example, a stimulus seen only in one eye is a stimulus to accommodation but not directly to vergence.)

When the vergence angle of the eyes is any value (v), vergence displacement ($\triangle v$) is defined as the angular deviation of the eyes from the natural vergence angle: $\triangle v = v - v_0$. Vergence-displacement energy is defined as a function of $\triangle v$: as $g_v(\triangle v)$, where $g_v(0) = 0$; that is, vergence-displacement energy is zero when the eyes are in the natural vergence angle. Any deviation ($\triangle v$) of the eyes from the natural vergence angle is assumed to have a larger vergence-displacement energy, according to a U-shaped function such as that illustrated in Figure 1.

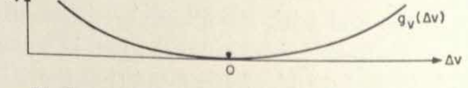


FIG. 1. Vergence-displacement energy, $g_v(\triangle v)$, as a function of vergence displacement, $\triangle v$. The marble representing vergence is at its equilibrium position, $\triangle v = 0$. Positive values of $\triangle v$ indicate convergence (relative to the equilibrium position); negative values indicate divergence. In this example, $g_v(\triangle v) = \log_e (1 + z) - z - .307$ where $z = [1 + (\triangle v)^2]^{1/2}$. This surface is the gravimetric analog of Hook's law: the horizontal restoring force (f) on the marble due to gravity is $f = -k\triangle v$.

The physical analogy here is a simplified version of the complete analogy, which is described below after all interactions have been considered. The curve $g_v(\triangle v)$ of Figure 1 represents a physical surface, a 'bowl.' A small marble rolls on this surface, and the projection of the marble's position on the horizontal axis represents the vergence angle of the eyes. The rule is that the marble tends to roll downhill, due to gravity. The force on the marble is proportional to the slope of the bowl ($-\partial e/\partial v$, Equations 1 and 4);[6] its motion depends on its mass and the friction it encounters (k_{1t} and k_{2t} respectively in Equation 4).

---

[5] Whenever accommodation strays from its nominal value, a negative-feedback optical device so alters the viewing distance of the objects that the eyes' normal attempt to maintain accurate accommodation then restores accommodation to the nominal value.

The g_v surface of Figure 1 represents the internal factors controlling vergence in the absence of an external stimulus to vergence. On this surface the marble rolls to a stable equilibrium position, to $\triangle v = 0$, the center of the g_v ($\triangle v$) curve. In order for the model to correspond to the eyes, the coefficient of mass of the marble is chosen sufficiently smaller than the coefficient of friction so that the marble rolls to the bottom without overshooting. In a real bowl, for example, this situation is achieved by filling the bowl with fluid. To move the marble from $\triangle v = 0$ requires 'work,' or 'energy.' The source of energy is image-disparity energy, described below.

Vergence image-disparity energy (h_v). Vergence image-disparity energy (h_v) represents the extent to which two retinal images agree. It is defined as follows. Let spherical coordinate systems for each retina be defined so that $x = y = 0$ is the fixation point (center of the visual field, line of sight) and so that corresponding points of the two retinas are assigned equal values of x and y. Let the illuminance distributions (images) on the left and right retinas respectively be $l(x, y)$ and $l'(x, y)$. The image-disparity energy $\triangle h_v$ in a small neighborhood ($\triangle x \triangle y$) of a point (x, y) is defined as

$$\triangle h_v(x, y) = |l(x, y) - l'(x, y)| \, \triangle x \triangle y. \qquad [5]$$

That is, $\triangle h$ is proportional to the absolute value of the amount by which l and l' differ.

The image-disparity energy (h_v) is given by summing $\triangle h_v(x, y)$ over the whole retina (R). In summing image disparity over the retina, central areas of the retina are given greater weight than peripheral areas. This selective weighting of the central area is accomplished by multiplying $\triangle h_v(x, y)$ by a weighting function $[w(x, y);\ 0 \leq w(x, y) \leq 1]$ before summing. Equation 6 then gives vergence image-disparity energy:

---

[6] Let $e' = \partial e(x)/\partial x$. In the system of Equations 1–4, the horizontal force (f) acting to change the vergence position (v) is defined by $f = -ke'$. In the gravimetric analog, $f_g = -k \sin e' \cos e' = -ke'/(1 + e'^2)$, which is an irrelevant complication. The gravimetric analog becomes equivalent to Equations 1–4 as $e'^2 \rightarrow 0$. Thus, by scaling units appropriately, the gravimetric analog satisfies the equations.

$$
\begin{aligned}
h_v &= \sum_R w(x, y) \triangle h(x, y) \\
&\approx \iint_R |l(x, y) - l'(x, y)|\, w(x, y)\, dx\, dy. \qquad [6]
\end{aligned}
$$

Equation 6 represents a point-by-point comparison of the left and right eye's retinal images (l and l'). When l is identical with l', h_v is zero. When l is not identical with l', h_v is greater than zero.

In practice, Equation 6 is somewhat inaccurate because it gives too much weight to large areas where l and l' differ slightly, and too little weight to small areas where l and l' differ sharply (boundary conflicts). This discrepancy arises because in the nervous system, each visual image is transformed before the two images are compared in the brain. The difficulty is overcome by not using $l(x, y)$ directly in Equation 6 but $T{:}l(x, y)$ where T represents the various retinal (and subsequent) neural transformations on l, such as adaptation and boundary enhancement. Another problem with Equation 6 is that the attention-weighting function $[w(x, y)]$ is not specified. The elucidation of neural transformations and of attention are interesting problems in visual psychology, but we can bypass them. The elementary definition of h_v by Equation 6 suffices for the present, providing that $w(x, y)$ is taken as any reasonable function that decreases as x and y increase.

Calculation of l. Given any external luminance distribution (L), we wish to calculate the retinal illuminance distribution (l). In general, L and L' (to left and right eyes respectively) could occur at any arbitrary distance from the observer and at any direction relative to his head; his eyes could be pointed in some other direction and be verged at another distance. Our interest here is in the effect of vergence upon retinal images, and we can avoid the hideous complications of this general formulation by restricting ourselves to symmetrical vergence; that is, to the case of an observer looking at stimuli directly in front of him. For this purpose it is convenient to define coordinates of the external stimuli (L and L') as those of the retinal stimuli they would produce if the observer's eyes were pointed straight ahead with a vergence angle of $v = 0$ (i.e., were verged at infinity.) Spherical coordinates are used, although so long as we restrict consideration to objects directly in front of the observer and to a small neighborhood of fixation, there is no important difference between Cartesian and spherical coordinates. The external horizontal coordinate (x) is given in angular units; $x = 0$ represents the projection at infinity of the midsaggital plane (the plane through the nose and perpendicular to the line connecting the centers of the two eyes). Positive values (angles) of x are to the right. The external vertical coordinate (y) is defined with positive values upward.

When $v = 0$, the retinal coordinates (x_r and y_r) and the external coordinates (x and y) coincide exactly, by definition. When the vergence angle is v, external and retinal coordinates are related by $x = x_r + v/2$ (left retina) and $x = x_r - v/2$ (right retina). For both eyes, it is assumed that $y = y_r$. For example, when the vergence angle is 2 deg., the left eye is rotated 1 deg. to the right (toward the nose) and the right eye is rotated 1 deg. to the left. The center of the left retina then is pointed at $x = v/2 = +1$ deg.

If we neglect the blur of the retinal image caused by the imperfect optics of the eye, then the retinal illuminance distributions (l and l') are exactly proportional to the external luminance distributions: $l(x_r, y) = kL(x, y)$ and $l'(x_r, y) = kL'(x, y)$, where k expresses the conversion factor from external luminance to retinal illuminance. Substituting the appropriate external coordinates for the left and right retinal coordinates gives simple expressions to relate l and L in terms of the eyes' vergence angle (v):

$$l(x - v/2, y) = kL(x, y) \quad \text{and} \quad l'(x + v/2, y) = kL'(x, y).$$

The expressions for l and l' contain the relevant information about the external stimulus, so that L and L' are not needed in subsequent calculations.

Calculation of $h_v(\triangle v)$. Let the external stimuli be $L(x, y)$ and $L'(x, y)$. Let the eyes be in their natural vergence position (v_0) for the state of accommodation, so that the retinal stimuli are $l(x - v_0/2, y)$ and $l'(x + v_0/2, y)$. The vergence image disparity (h_v) for this pair of retinal images is given by substituting l and l' into Equation 6. Now, let the eyes alter their vergence by $\triangle v$, to $v = v_0 + \triangle v$. The vergence image disparity still is given by substituting l and l' in Equation 6; now it is a function of $\triangle v$, namely,

$$h_v(\triangle v) = \iint_R |l[x - (v_0 + \triangle v)/2, y] - l'[x + (v_0 + \triangle v)/2, y]|\, w(x, y)\, dx\, dy. \qquad [7]$$

Equation 7 is at the very heart of the theory so it is important that the reader understand it completely. To help intuitive understanding of Equation 7, the concept of a cyclopean image is introduced here. Our two eyes receive different retinal images (l and l') of the world; yet, we are unaware of having two separate images and perceive things as single, like the one-eyed Cyclops. For the moment we define the cyclopean image ($^*l$) simply as the sum of the two separate images:

$$^*l(x, y) = l(x, y) + l'(x, y).$$

(Later, in the section on the neural model and its predictions for binocular rivalry, we shall see that the subjective cyclopean view is a much more complicated combination of l and l' than their sum.)

When the external stimuli (L and L') remain fixed, then the retinal images (l and l', and hence $^*l$) depend on the vergence angle of the eyes. Figure 2c illustrates a pair of external stimuli (L and L'); the cyclopean images that would result for two different vergence positions are shown in Figure 2ab.

The critical term of Equation 7 is $|l - l'|$, which is like a cyclopean image except that l and l' are subtracted instead of being summed. Since image-disparity energy (h_v) depends on the difference between l and l', it is more nearly zero the more nearly identical l and l' are. For the vergence position v_0 of the eyes illustrated in Figure 2a, $h_v(v_0)$ is at a relative minimum because the two rectangles are optimally superposed. The image-disparity energy $h_v(v_0)$ is not zero, however, because neither the rectangles nor the disks superpose exactly, and thus they fail to cancel completely. When the vergence angle of the eyes is v_1, the two disks are superposed but the rectangles fail to superpose. Image-disparity energy calculated at $\triangle v_1$ again is at a relative minimum, but $h_v(\triangle v_1) > h_v(0)$ because of the large contribution of the now unsuperposed rectangles.

Figure 2 also illustrates that changing vergence of the eyes is equivalent to keeping the eyes stationary but shifting the positions of the physical stimuli (L and L'). The image-disparity energy function $h_v(\triangle v)$ is an index of the similarity of L and L' as a function of their relative position. However, Equation 7 is written in terms of the retinal stimuli (l and l') in order to emphasize that the eyes need not actually verge to compute $h_v(\triangle v)$. A single pair of retinal images (l and l'), obtained in a single glimpse, contains

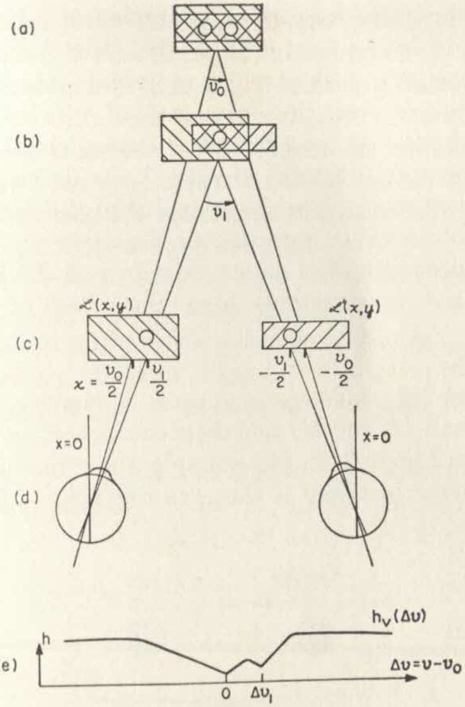


FIG. 2. Calculation of image-disparity energy, $h_v(\triangle v)$. (d) The eyes are in the natural vergence position, v_0, induced by accommodation on the retinal stimuli, $L(x, y)$ and $L'(x, y)$, illustrated in (c); the vertical position of L and L' is arbitrary. (a) Cyclopean image illustrating $l(x - v_0/2,y) + l'(x + v_0/2,y)$, from which $h_v(0)$ is calculated. (b) Cyclopean image illustrating $l[x - (v_0 + \triangle v_1)/2,y] + l'[x + (v_0 + \triangle v_1)/2,y]$, from which $h_v(\triangle v_1)$ is calculated. (e) The $h_v(\triangle v)$ for the image pair (c), as given by Equation 7. The abscissa, $\triangle v$, is in angular units of vergence. Clockwise rotations of the left eye and counterclockwise rotations of the right eye are indicated by positive values of $\triangle v$, which represent convergence.

all the information needed to compute the entire function $h_v(\triangle v)$. In other words, a single glimpse suffices to 'tell' the visual system the cyclopean image it should expect to see in any vergence position, and from this information it can compute $h_v(\triangle v)$. Just how much of this potentially available information actually is used by the visual system is considered below. Here, for convenience, the model is stated as though the entire function $h_v(\triangle v)$ were known; in fact, knowledge of just two values of $h_v(\triangle v)$ at any one time would be a sufficient assumption.

Figure 2e illustrates $h_v(\triangle v)$ as a function of $\triangle v$. Two interesting properties of $h_v(\triangle v)$ are evident: first, that $h_v(\triangle v)$ consists of the sum of functions, each of which represents registration of some part of the images and, therefore, each of which is symmetrical about its minimum (if the effect of w is negligible); and, second, that at any relative minimum of $h_v(\triangle_v)$, its derivative $dh_v(\triangle v)/d(\triangle v)$ is discontinuous. The discontinuity of the derivative results from the absolute-value operation in Equation 3. As a practical matter, such discontinuities do not exist in real life because image blur and 'noise' (uncertainty) have the effect of rounding the minimum.[7]

Examples of the model for vergence. A simpler illustration of binocular stimuli (L and L') and their corresponding $h_v(\triangle v)$ function is given in Figure 3. In this example, the retinal images (l and l') would become identical if the eyes diverged slightly. Thus in

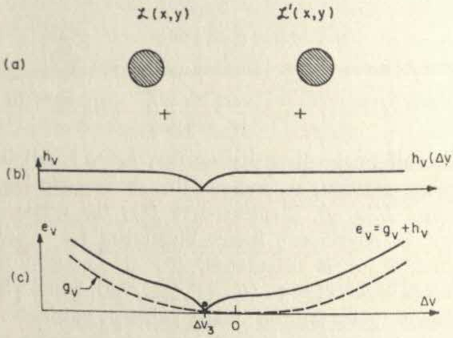


FIG. 3. A visual display with a single stable vergence state. (a) The left eye sees $L(x, y)$, the right eye $L'(x, y)$; the crosses in L and L' indicate the point $x = y = 0$; the disks are symmetrically displaced from $x = y = 0$ (nasally on the retinas, temporally in the external world). (b) Vergence image-disparity energy, $h_v(\triangle v)$, as a function of vergence displacement, $\triangle v$, for the display illustrated in (a). (c) Vergence energy, $e_v(\triangle v)$, as a function of $\triangle v$. The point $\triangle v_3$ indicates the projection on the horizontal axis of the minimum of the energy well in e_v. The marble representing vergence is in the energy well, its only stable position.

---

[7] The discontinuity of its derivative at zero gives the absolute-value function some rather strange properties that probably do not accurately reflect nature (which abhors discontinuous derivatives). Taking into account the effect of the neural transformation on l and l' before $h_v(\triangle v)$ is computed improves matters (see Appendix A).

Figure 3b, $h_v(\triangle v_3) = 0$. Figure 3c illustrates vergence energy $e_v(\triangle v)$ defined as the sum of $g_v(\triangle v)$ and $h_v(\triangle v)$.

The binocular stimuli (L and L') of Figure 3a yield an $e_v(\Delta v)$ function that has only one relative minimum. For this function, there is one and only one stable position of the marble, namely, at the bottom of the $e_v(\triangle v)$ function. This means that no matter what the initial vergence position of the eyes may be, they ultimately stop at the vergence position $\triangle v_3$.

Consider now what happens when L and L' of Figure 3a are separated horizontally; that is, when the left eye's stimulus is moved to the left and the right eye's stimulus to the right. Vergence becomes increasingly difficult. Figure 4 illustrates that a marble initially at rest (at $\triangle v = 0$) does not roll into the energy well at $\triangle v_4$ but remains at $\triangle v = 0$. However, if the left and right eyes are initially verged on their stimulus at a displacement of about $\triangle v_3$ (Figure 3), and then the two stimuli are slowly separated, the marble remains in the energy well. The eyes remain verged. This example shows that the stimuli illustrated in Figure 4 permit two stable states of vergence.

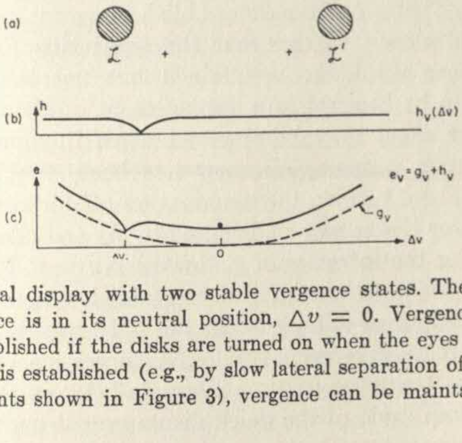


FIG. 4. A visual display with two stable vergence states. The marble representing vergence is in its neutral position, $\triangle v = 0$. Vergence on the disks cannot be established if the disks are turned on when the eyes are at $\triangle v = 0$. Once vergence is established (e.g., by slow lateral separation of the disks from the displacements shown in Figure 3), vergence can be maintained at $\triangle v_4$.

When L and L' are separated by the amount illustrated in Figure 5, the energy well in the $e_v(\triangle v)$ curve has flattened out to such an extent that it no longer can hold the marble, which then rolls to the position of zero vergence displacement. The binocular stimulus pair of Figure 5 thus yields only one stable vergence position for the eyes.

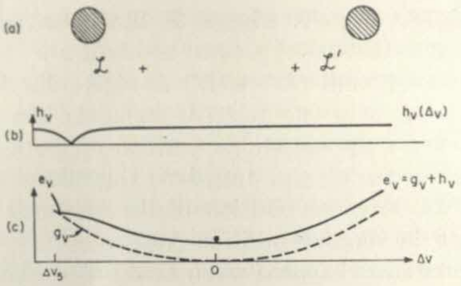


FIG. 5. A display for which vergence cannot be established. The disks are displaced laterally even further than those illustrated in Figure 4a. The marble representing vergence is in its only stable position, $\triangle v = 0$.

Suppose that L and L' initially are separated by the amount illustrated in Figure 5. Then, to establish vergence they would have to be brought closer together than the separation $\triangle v_4$ (Figure 4), where vergence would be maintained but not established; they would have to be brought to a displacement of about $\triangle v_3$ (Figure 3). A similar effect may be observed when illumination is interrupted. Suppose vergence somehow has been established at $\triangle v_4$ (Figure 4). Then, turning the illumination off for a few tenths of a second removes the energy well caused by h_v and allows the marble to move under the influence of g_v toward $\triangle v = 0$. It then requires a slow separation of the stimuli to reestablish vergence at $\triangle v_4$.

Slow separation of stimuli is not the only way to establish vergence at extreme divergence values. A succession of small, quick displacements (with longer time intervals between them) also succeeds. The amplitude of the quick displacement must be such that the marble remains inside the energy well. Then, during the interval between displacements, it rolls down into the new well. The process may be repeated until the critical divergence illustrated in Figure 5 is reached.

'Vergence disparity.' When the minima of $h_v(\triangle v)$ are rounded at their bottom [i.e., $dh_v(\triangle v)/d(\triangle v)$ is continuous at a minimum], then the minima of $e_v(\triangle v)$ do not correspond exactly to the minima of h_v but are shifted toward the center of e_v, toward $\triangle v = 0$. The difference in location of minima is negligible when they lie near $\triangle v = 0$, but it becomes significant as $\triangle v$ increases, because the e_v function becomes steeper.

In the model, the marble rests only at minima of the e_v function.

When a minimum of e_v does not coincide exactly with a minimum of $h_v(\triangle v)$, it means that the eyes verge insufficiently to carry their retinal stimuli to exactly corresponding points. This phenomenon of human vision is known as vergence disparity and has been measured under a variety of conditions.[8] The psychophysical measurements of vergence disparity are qualitatively in agreement with those predicted by the model. Further comparisons between the model and experiment are deferred until the full interactions between vergence and accommodation have been considered.

Model of Vertical Vergence and of Torsional Rotation

The vergence model can be applied to two other kinds of eye movements. Vertical vergence occurs when the stimulus to one eye is displaced vertically. The eye can move vertically to again center itself on the displaced stimulus, even while the other eye remains relatively stationary. Similarly, when the stimulus to one eye is rotated around the line of sight, the eye can rotate to realign itself with the stimulus. These movements occur only within certain bounds, about 6 deg. for vertical vergence and about 7 deg. for torsional rotation (see n. 3, pp. 58–62).

The angle between the vertical elevation of the two eyes is exactly analogous to the horizontal vergence angle, and the same model applies to control of both kinds of angle. In the case of torsion, the torsional angle of the eye is the analog of the horizontal vergence angle, and the model applies independently to each eye. In fact, in composing demonstrations of multistability phenomena, it often is more convenient to use vertical vergence or torsional rotation, because these movements are less under voluntary control than is horizontal vergence.

Model of Accommodation

Just as the current state of accommodation determines the natural vergence position, the value of vergence determines the natural accommodation position. To displace accommodation from its natural position—from its value in the absence of any stimulus to accommodation—requires an amount of energy described by the accommodation-displacement energy function (g_a); this energy is supplied by a stimulus to accommodation via an image-defocus function (h_a); and the actual state of accommodation is predicted from the accommodation energy function (e_a): $e_a = g_a + h_a$. The model for accommodation is thus formally identical to the model for vergence and differs only in the details of calculating the functions. These details offer some special problems, which are considered below.

---

[8] Among others, see K. N. Ogle, *Researches in Binocular Vision,* 1950; E. F. Fincham and J. Walton, The reciprocal actions of accommodation and convergence, *J. Physiol. (London),* 137, 1957, 488–508.

Units of accommodation. The conventional measurement unit of accommodation is the *diopter*. When the eye is focused on an object $1/n$ meters distant from it, it is said to be accommodated to n diopters. The scales of diopters of accommodation and of degrees of vergence are related almost linearly. For objects in front of the observer, the deviation from linearity is of the third order and does not exceed about 1% at 8 diopters of accommodation, which is the practical limit of the adult range. Therefore with negligible error, we can use the same scale for accommodation as for vergence. There is, in fact, an advantage in using the diopter scale: it is independent of the interpupillary distance (p) of the observer. For an average value of p (6.0 cm.), 1 diopter = 3.5 deg.

Accommodation-displacement energy (g_a). Theoretically, accommodation-displacement energy (g_a) may be regarded as a function of only Δa, the displacement of accommodation from the value (a_0) determined by vergence. In practice, however, the value of accommodation is strongly influenced by the mechanical properties of the lens.[9] The situation may be conceived of as one in which neural signals attempt to control accommodation according to a function $[g_a(\Delta a)]$ that depends only on Δa; but the transduction of these signals into accommodative power also depends on a. The nonlinear (a-dependent) transduction of attempted accommodation into achieved accommodation is an important and unfortunate consequence of the hardening of the lens caused by aging.[10] It is worth noting that rotation of the eyeball in vergence also is limited by mechanical factors; in vergence, however, the neural control operates so far within the mechanical limits that the mechanical factors can be neglected.

---

[9] For a summary of the evidence for distinguishing between the muscular force exerted on the lens and its mechanical response, see M. Alpern, W. M. Kincaid, and M. J. Lubeck, Vergence and accommodation: III, Proposed definitions of the AC/A ratios, *Amer. J. Ophthal.*, 48, 1959, 141–148.

[10] See Fincham and Walton (n. 8 above).

Image-defocus energy (h_a). Image-defocus energy (h_a) represents the blurredness of a retinal image. Although we do not know what particular h_a is computed by the nervous system, we shall define an h_a with these three properties: it is a function that *could* be computed by neurons; it is monotonically related to blur; and it is mathematically tractable. Given $l(x, y)$ and $l'(x, y)$, the illuminance distributions on the retinas, an h_a satisfying the three conditions is defined by

$$-h_a = \iint_R [\nabla^n l(x, y)]^2 w(x, y)\, dx\, dy$$

$$+ \iint_R [\nabla^2 l'(x, y)]^2 w(x, y)\, dx\, dy. \qquad [8]$$

In words, $-h_a$ is the square of the magnitude of the Laplacian of $l(x,y)$, weighted by $w(x,y)$, integrated over the retina, and summed for the two eyes.

The algebraic value of h_a as defined by Equation 8 increases when $l(x,y)$ is blurred and decreases when $l(x,y)$ is sharpened. Specifically, if $l(x, y)$ represents any luminance distribution on the retina and if $l_b(x, y)$ represents this distribution after it has been blurred, it is shown in Appendix A that

$$\iint_R [\nabla^2 l(x, y)]^2\, dx\, dy \geq \iint_R [\nabla^n l_b(x, y)]^2\, dx\, dy, \qquad [9]$$

where ∇^n represents any order of differentiation or integration applied to l.

The reason for choosing the Laplacian—rather than $l^2(x, y)$ directly, for example—is that it represents a good approximation to the spatial analysis performed by the retinal receptive fields (see Appendix A). The reason for using a square inside the integral of Equation 8 instead of the absolute value (as in the definition of h_v) is for the mathematical elegance of the theorems which follow therefrom. In fact, the same general properties hold for the absolute value as for the square (Appendix A).

Calculating $\partial e_a/\partial a$; How the visual system estimates h_a. To compute h_a as a function of Δa, the accommodation displacement, it need only be noted that l is implicitly a function of a [i.e., $l = l(x, y, a)$]. Substitution in Equation 8 gives $h_a(\triangle a)$. This, however, leads to a problem.

In the model, $\partial e_a/\partial a$ controls the behavior of the marble. To know $\partial e_a/\partial a$ it is necessary to know $\partial h_a/\partial a$, and to estimate $\partial h_a/\partial a$ requires knowledge of at least two values of h_a. The problem is one of determining how the nervous system computes h_a for more than one value of a when the lens can assume only one value of a at any one time, for any one wavelength of light at any one point. These restrictions suggest three possible solutions: that the visual system obtains two (or more) values of h_a by determining a at two or more times, by determining a at two or more wavelengths, or by determing a at two or more points. Let us examine each of these solutions in turn.

The temporal solution is a trial-and-error solution. When the eye is confronted with a new stimulus, it would alter its present value of accommodation (a_1) to a new value (a_2, a trial value). The new value (a_2) might be an improvement or not, but no matter which, it would provide a second value of h_a, from which $\partial h_a/\partial a$ may be estimated by

$$\partial h_a/\partial a = [h_a(a_2) - h_a(a_1)]/(a_2 - a_1).$$

In such temporal trial and error, the initial change of accommodation in response to a new stimulus would be in error about half the time; subsequent changes of accommodation would be in the right direction. However, the first measurable accommodation response to a step or to a pulse displacement in the accommodative distance of a stimulus is generally in the right direction.[11] We must infer that at least two values of h_a are computed before accommodation has changed.

The chromatic aberration of the lens, which introduces a $\triangle v$ of about 1.3 diopters ($\triangle v \approx 4.5$ deg.) between the long and short wavelengths, provides a second possible way of realizing simultaneously two values of a on the retina. This cannot be the only mechanism, because some subjects make initially correct accommodation responses even in monochromatic illumination.[12]

---

[11] F. W. Campbell and G. Westheimer, Dynamics of accommodation responses of the human eye, *J. Physiol.*, 151, 1960, 285–295; E. F. Fincham, The accommodation reflex and its stimulus, *Brit. J. Ophthal.*, 35, 1951, 381–393.

[12] Fincham (n. 11).

The alternative to time and wavelength as means of obtaining different values of a is spatial variation of a. A mechanism for spatial variation of a is known to exist. In the neighborhood of the optic axis, the retina is not exactly perpendicular to the optic axis; thus, some parts of the retina are optically closer to the lens than others.[13] Comparison of two regions in different focus would provide an estimate of $\partial e_a/\partial a$ to even the stationary eye. By elimination, then, spatial variation of a is the means by which the relevant position of $e_a(a)$ is computed by the visual system. This complication requires: redefining $l(x, y, a)$ as $l[x, y, a(x, y)]$; calculating the predicted difference $\triangle h/\triangle a$ (rather than $\partial h/\partial a$) from Equation 8; and using two different weighting functions for the two terms of the difference [i.e., $\triangle h = h(w_1, l, l') - h(w_2, l, l')$]. These complications, though real, are not of great theoretical importance for the understanding of accommodation.

Whether h_a is known simultaneously for only two values of $\triangle a$ or for infinitely many values makes little difference. Only two values are needed to estimate $\partial e_a/\partial a$, and this estimate suffices to control the dynamics of a. However, when temporal information is used to estimate $\partial e_a/\partial a$ (and it occasionally may be), the situation can be vastly more complicated. For example, if it is assumed that there is a fixed time interval ($\triangle t$) between the two instants of sampling h_a, then the estimate depends on $\triangle a/\triangle t$, the average velocity of a during $\triangle t$. This leads to an additional term $[k_3(\partial e_a/\partial a)(-\partial a/\partial t]$ on the right-hand side of Equation 4 and converts a highly stable system into a highly unstable one, depending on how large k_3 is. Such a system would tend to have big, slow drifts in accommodation because when da/dt is small, the stabilizing force $\partial e_a/\partial a$ is small. Even small amounts of extraneous 'noise' would perturb a far from the minimum of e_a. In fact, this kind of 'd.c. drift' is a distinguishing characteristic of accommodation.

Demonstration of the multistability of accommodation. If the reader is young enough to possess a lens still capable of a few diopters of accommodation, he can easily demonstrate to himself two different stable states of accommodation in response to a single external stimulus configuration. First, cover one eye and establish the near point for accommodation by moving a pencil toward the open eye, while focusing on its point. (The near point is the smallest distance at which a sharp image is possible.) Second, fixate a textured object (say, a picture) a few meters away. Third, while maintaining focus on the picture, move the pencil to the near point. Move it to various positions in front of the picture so that it obscures the point of fixation. Maintain focus on the picture. Fourth, focus on the pencil keeping the picture in the background. Then, alternate the third and fourth steps until you find a particular line of sight for the eye and a particular position of the pencil such that they on some occasions lead to accommodation of the picture and on other occasions to accommodation of the pencil. It is easy to focus either on the picture or on the pencil, but these changes of focus usually are consequences of changes in line of sight. The purpose of alternating the third and fourth steps is to find a single external stimulus configuration (same line of sight) that can yield either state of accommodation. The external stimulus (picture plus pencil) thus induces two different states of accommodation (accommodation on picture, accommodation on pencil); there are two stable response states to one stimulus (multistability).

---

[13] G. L. Walls, *The Vertebrate Eye and Its Adaptive Radiation*, 1942, 7 and 255.

In this demonstration, vergence of the covered eye changes with accommodation. The correlated vergence change is irrelevant to the demonstration; it means only that the demonstration may not reveal the dynamic parameters of accommodation when the covered eye exerts 'drag' on the changing value of accommodation.

Model of Fusion

When the eyes are at a given vergence angle (v), this vergence position defines a surface called the horopter. The defining characteristic of the horopter is that whenever a point lies on the horopter, it stimulates corresponding points of the two retinas.[14] In a small neighborhood of fixation, the horopter may be considered a plane, which is called the vergence plane, or plane of fixation. The horopter corresponding to v defines the optimum depth of an object for fusion of the object. When an object lies slightly in front of or behind the horopter, fusion is more difficult. When such objects are fused, they are seen in depth relative to objects on the horopter. This depth sensation is here called *fusion depth.*[15]

---

[14] It is assumed here that the nonius method is used for the empirical determination of the horopter, as suggested by Ogle (n. 8). Also see K. N. Ogle, The problem of the horopter, in H. Davson (ed.), *The Eye*, 4, 1962, 325–348. By this method, there is no significant difference in the neighborhood of fixation between corresponding points (horopter) and corresponding coordinates (as defined early in this paper).

The model of fusion is formally equivalent to the models for vergence and for accommodation. First, in the fusion model the fusion distance (distance from eyes to plane of fusion) corresponds precisely to the vergence distance in the vergence model (distance from eyes to plane of vergence). Second, the perceived fusion depth in the fusion model corresponds to the achieved vergence angle in the vergence model. Third, the calculation of fusion disparity from the retinal stimulus (l and l') is the same as the calculation of vergence disparity, except that fusion is considered separately for each small retinal neighborhood (in vergence, the relevant region is the whole retina). Fourth, in the fusion model there is an internal bias to fuse objects in the plane of vergence (the corresponding bias in the vergence model is the bias to verge on the plane of accommodation). Fifth, the external stimulus energies and internal bias energies add to form the surface governing the fusion-depth plane just as the corresponding vergence energies add to form the e_v surface that determines the vergence plane.

Before developing the model for fusion in detail, it must be noted that when an object lies too far behind or in front of the horopter, fusion is impossible. Nevertheless, the object still may seem to be located at the appropriate depth plane. This depth sensation is here called *postfusion depth;* it corresponds to what is conventionally called 'qualitative stereopsis.' In the physical model of fusion phenomena, the perception of fusion and the perceived fusion depth are two different manifestations of the same underlying variable: the plane of fusion. By taking fusion depth and postfusion depth to be two different regions of a single continuum of perceived depth, the physical model for fusion depth (which extends over about $\pm\frac{1}{8}$ deg., or ±.03 diopters, in the fovea) may be extended to include postfusion depth (about ±2 deg. in central vision) without introducing any new principles. The complete, complex relations between perceived fusion, fusion depth, and postfusion depth are described in the neural model. Here we consider only the most basic phenomena and how the model applies to them.

---

[15] 'Fusion depth' corresponds approximately to what has been called 'patent stereopsis'; see K. N. Ogle, On the limits of stereoscopic vision, *J. exp. Psychol.*, 44, 1952, 253–259. Because of binocular rivalry, however, the range of patent stereopsis may vastly exceed the range of fusion depth—see the sections on binocular combinations.

Units of fusion displacement ($\triangle u$). The unit of vergence was defined as the angle between the lines of sight of the eyes. The vergence angle is also an unconventional measure of distance from which conventional distance can be calculated. The distance (d) from the midpoint between the observer's eyes to an object located directly in front of him is given by

$$d = \left(2 \tan \frac{u}{2}\right)^{-1} p \xrightarrow[u \to 0]{} p/u, \qquad [10]$$

where p is the interocular distance and u is the vergence angle of the object. Although d, the linear distance from the observer to the plane of fusion, is the conventional unit of distance, it is more convenient to designate the plane of fusion by u, its *angular distance* as defined by Equation 10. This scale of distance for u is commensurate with the scales for v and a, and it shares their advantages of being directly proportional to distance on the retina (see Figure 2d). When u is small, the tangent function can be removed from Equation 10, giving directly the reciprocal 'diopter' relation between d and u. In considering fusion, the distance of the horopter is v, the distance of the plane of fusion is u, and the distance from plane of fusion to the horopter is $\triangle u = u - v$. This follows the earlier sign convention for $\triangle v$; points beyond the horopter are represented by $\triangle v < 0$.

According to the model, whether a pair of retinal images (l and l'; an object) is seen as fused or unfused is determined by the fusion energy function $e_u(\triangle u)$ corresponding to that object, in the same general way that the vergence energy function $e_v(\triangle v)$ determined whether the eyes verged or did not verge on the object. The $\triangle u$ value of an e_u minimum determines the perceived fusion depth. Two components sum to give fusion energy: fusion-displacement energy (g_u) and image-disparity energy (h_u).

Fusion-displacement energy (g_u). Fusion-displacement energy (g_u) is given by a U-shaped function [$g_u(\triangle u)$], which governs whether or not fusion occurs. The function does not extend more than about $\frac{1}{8}$ deg. to either side, representing the fact that fusion is impossible for objects displaced more than about $\frac{1}{8}$ deg. from the horopter. The function $g_u(\triangle u)$ may also be interpreted as governing the depth sensation that results from successful fusion. In this case, it is useful to think of it as having an extension to ± 2 deg. as fusion depth (± $\frac{1}{8}$ deg.) and postfusion depth (± 2 deg.) form

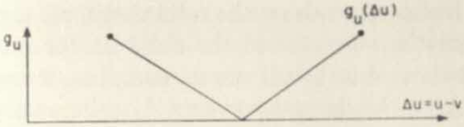


Fig. 6. Fusion-displacement energy, g_u, as a function of fusion displacement, $\triangle u$. The function $g_u(\triangle u)$ is defined only for $|\triangle u| < \frac{1}{8}$ deg. in central vision. The same function extended to $|\triangle u| < 2$ deg. describes the corresponding continuum of perceived depth (fusion depth plus postfusion depth).

a single continuum. The sharp minimum of $g_u(\triangle u)$ represents the fact that unfused objects (objects seen by only one eye) tend to be seen in the plane of the horopter; the sides of $g_u(\triangle u)$ are drawn straight (see Figure 6) mainly to help the reader distinguish g_u from g_v on graphs where both functions occur.

Fusion image-disparity energy (h_u). The contribution of h_u to e_u is defined as was the contribution of h_v to e_v, with one difference: whereas h_v was calculated by summing disparity over the whole retina, h_u is calculated in each small neighborhood $(\triangle z)$ of the retina and summed over only those neighborhoods which cover the object. When several objects are present simultaneously in the visual field, $h_u(u, z)$ is calculated separately for each of them.

Examples of the model for fusion. The study of fusion in isolation requires the control of $\triangle u$, which in turn requires control of vergence movements. Some voluntary control of vergence is possible with the aid of a strong vergence stimulus (e.g., a picture frame within which other stimuli are presented). A better method is that of Fender and Julesz (n. 4), who studied the fusion properties of stimuli whose retinal position was stabilized by optical means (i.e., stimuli whose retinal position was independent of vergence). In general, they demonstrated that fusion had the same characteristics as vergence, although they did not analyze their data in this way. Their results can be summarized efficiently in the terminology of the model.

Fender and Julesz observed that a stabilized image pair could always be seen as a single object if the minimum of their $h_u(\triangle u)$ function occurred for $\triangle u < \frac{1}{2}$ deg. (compare with Figure 3) and could conditionally be seen as a single object (by slow separation of the images) for $\triangle u$ in the range $\frac{1}{2}$ deg. $< \triangle u < 1$ deg. (compare with Figure 4). Their results show, as the model predicts, that the range of the fusion depends on the retinal stimuli (l and l') used—particularly on the steepness of the sides of the minimum of the $h_u(\triangle u)$ function—since their most complex textured stimulus, which would give the deepest, steepest h_u minimum, also gave the greatest range of single vision. (The quantitative discrepancy between the range of single vision observed by Fender and Julesz and the range of fusion predicted by the model is explained in the section on predictions of the neural model, below.)

One interesting incidental observation of Fender and Julesz is the existence of various stable *'secondary' fusion points.* This too is predicted by the model. The secondary fusion points occur because a complex retinal stimulus (l and l') may yield an $h_u(\triangle u)$ function that has several smaller relative minima in addition to its major minimum. The minor minima result from chance superposition of unrelated but similar parts of the complex images. Such minor minima provide alternative fusion modes. In normal vision, saccadic eye movements produce a 'noise' perturbation of vergence $(\triangle v)$. In normal vision, therefore, fluctuations in vergence from one eye movement to the next make it unlikely that the point of fusion could long remain in a small, shallow relative minimum, especially when it is near a major minimum of the $e_u(\triangle u)$ function. In stabilized vision, the eye movements are optically canceled and therefore secondary fusion points are more easily observed.

Panum's fusional area. In the case of vergence movements, 'vergence disparity' occurred because the minima of the $e_v(\triangle v)$ function did not coincide exactly with the minima of the $h_v(\triangle v)$ function. The same displacement phenomenon occurs in the fusion model when the minima of $e_u(\triangle u)$ function fail to coincide exactly with the minima of the $h_u(\triangle u)$ function. This means that when retinal fusion occurs for a stimulus that is not in the horopter, the apparent plane of the fused object is shifted slightly from the true stimulus plane and toward the horopter. The more important psychophysical correlate—insofar as the shift is detectable—would be blurring and then doubleness of vision of the object. Blurring might be observed even when the error in fusion depth itself is too slight to be measured.

The retinal range within which fusion is possible without noticeable blurring or double images is called Panum's fusional area. It traditionally has been measured with fine-line stimuli and, so measured, is remarkably small, about $\frac{1}{10}$ deg. (6 min. of arc) in foveal vision.[16] We now examine the predictions of the model of the range of singleness of vision (Panum's area), of the range of fusion (which is larger than Panum's area), and of how these ranges depend on the nature of the stimuli used to measure them.

In the model, the correlate of doubleness of vision is the shift of a minimum of h_u caused by adding g_u to form e_u. (The analogous effect in the model of vergence produced 'vergence disparity.') We can make a quantitative statement. The displacement of the minimum of h_u is small when it is a sharp minimum (i.e., when the sides of h_u near its minimum are steep). This sharpness is determined by d^2h_u/du^2; the greater d^2h_u/du^2, the greater the range of singleness of vision. The range of fusion is determined by the maximum slope of h_u (max dh_u/du); the greater this quantity, the greater the range of fusion (see Figure 5).

We can say immediately that fine lines are poor stimuli for vergence and for fusion because the minimum of h_u is very shallow; large-area, fine-textured, high-contrast stimuli produce the deepest minima in h_u, and thereby the largest values of d^2h_u/du^2 and max dh_u/du. Therefore the physical model correctly predicts that large-area stimuli produce a greater range of singleness of vision and a greater range of fusion than do fine lines. It also predicts that the range of fusion (determined by max dh_u/du) may far exceed the range of singleness of vision (determined by d^2h_u/du^2). This prediction accounts in part for the paradox of the enormously large range of binocular fusion in casual observations (several degrees) and the tiny range of Panum's area ($\frac{1}{10}$ deg.). The difference between the fine-line stimuli used to measure Panum's area and the large-area stimuli encountered in normal vision is not the whole explanation, however, because binocular rivalry is involved. The problem is considered again in the section on predictions of the neural model.

THE PHYSICAL MODEL—WITH INTERACTIONS

Up to this point, the physical model has dealt with each system in isolation. The model for vergence considered a and u fixed; the model for accommodation considered v and u fixed; and the model for fusion that was just discussed considered v and a fixed. We are now ready to consider the interactions between the three mechanisms and to describe the operation of the complete model.

---

[16] Ogle (n. 8).

Interaction of Vergence and Accommodation

In Equations 1 and 2, both v and a occur in both terms on the right side of the equations: the displacement terms (g) contain a and v explicitly, and the image terms (h) contain a and v implicitly in l and l'.

The image-term interactions are relatively unimportant and may merely be mentioned. In Equation 2, vergence affects the position on the retina of the stimulus (l and l') to accommodation. Because the weighting function $[w(x, y)]$ is not a constant, retinal positions differ in the effectiveness with which they affect accommodation. Because vergence can shift l and l' relative to w, it can affect h_a slightly. Similarly, in Equation 8, the value of accommodation affects the shape of minima in the h_v function; blurred images have shallower minima than sharp images. In practice both of the effects of these image-term interactions are negligible compared to the effects of the displacement-term interactions.

The displacement-term interactions between vergence and accommodation are evident in Equations 1 and 2. In the accommodation-displacement term of Equation 2 $[g_a(\triangle a)]$, $\triangle a$ is defined in terms of both a and v, by $\triangle a = a - v$. Thus, the value of v determines the center of the accommodation-displacement function. Similarly, in Equation 1, the center of the vergence-displacement term g_v is determined by the value of a. These relations are incorporated into the model illustrated in Figure 7. This model consists of two templates (g_a and g_v), which slide sideways, parallel to each other, along rails to which they are attached. Two marbles (v and a, representing vergence and accommodation respectively) roll along the templates. A constant force (a spring, not gravity, in this model) drives each marble toward the 'bottom' of each template (i.e., toward the opposite rail). The v marble is held by a rod-and-spring device which is attached rigidly and perpendicularly to the center of the g_a template, so that when the marble moves, the g_a template slides horizontally by exactly the same amount. The a marble is similarly attached to the g_v template. Figure 7a shows the systems temporarily out of equilibrium, at the instant when a visual stimulus has been turned off; Figure 7b illustrates the equilibrium reached in the absence of a stimulus.

Figure 8ab illustrates the vergence and accommodation systems' response to a stimulus to convergence. In Figure 8a the stimulus has just been turned on and the systems are starting from their previous

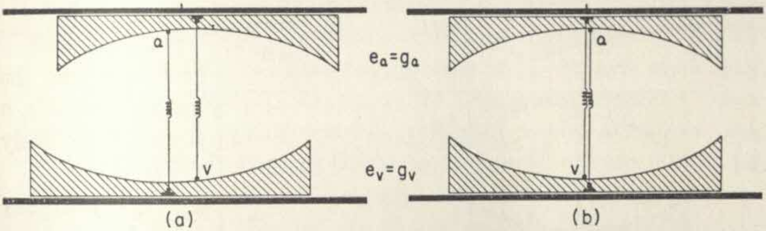


FIG. 7. Vergence-accommodation interaction in the dark. The marble representing vergence rolls on the g_v template, pushed toward the lower rail by a constant force from the spring; the g_a template, rigidly attached to the rod-and-spring device holding the v marble, follows the v marble exactly. The accommodation marble rolls on the g_a surface, pushed toward the upper rail by its spring; it carries the g_v template with it. The templates slide along the rails without friction. (a) The v-a systems temporarily out of equilibrium; (b) the final equilibrium. No visual stimulus to vergence or to accommodation is present here.

equilibrium position. Figure 8b illustrates the equilibrium position reached in response to this new stimulus. The image-displacement functions h_a and h_v are also illustrated in Figure 8. They remain fixed in the external world as the g_a and g_v templates slide. The e_a and e_v surfaces upon which the marbles roll are formed by adding the stationary hs to the sliding gs. The physical realization of the $g + h$ addition is a cumbersome, irrelevent mechanical problem; it is trivial in an electrostatic analogy.

Figure 8 also illustrates how accommodation helps vergence. If accommodation were fixed at the initial value, the vergence energy function e_v of Figure 8a would indicate a bistable vergence configuration. Starting from the position of Figure 8a, vergence on the stimulus would not occur (i.e., the marble would not leave its initial position). However, when accommodation is free to vary, it carries the vergence template (g_v) along with it until vergence also occurs.

Figure 8cd illustrates what happens when the stimuli to accomodation and vergence do not agree. This would happen, for example, if a person with normal vision, who had verged and accommodated on a stimulus, suddenly put on reading spectables with, say, a correction of +1 diopter ($\approx$3.5 deg.). Figure 8c illustrates the situation the instant the spectacles have been put on. His vergence is correct, but his accommodation is now too large by +1 diopter and, as Figure 8c illustrates, the a marble would start to roll into the e_a energy well. As it moves, it carries with it the g_v template. There

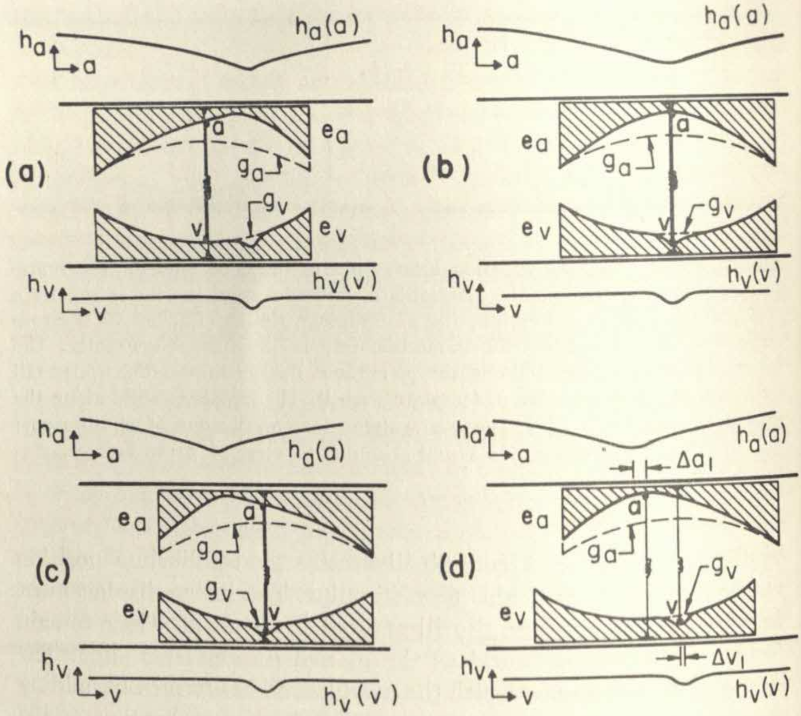


FIG. 8. Vergence-accommodation interaction. Vergence image-disparity energy, h_v, is shown below the vergence template's rail, and accommodation image-defocus energy, h_a, is shown above accommodation's rail. The image functions, h_a and h_v, add directly to the displacement functions, g_a and g_v, to form their respective energy functions, e_a and e_v (Equations 1 and 2). Here, g_a and e_a (but not h_a) are represented 'upside down.' The a and v marbles roll on e_a and e_v surfaces propelled outward by their springs; they carry the opposite templates with them, and in so doing cause the e_v and e_a surfaces to change shape continuously as g and h are added in different spatial relations. (a) The v-a systems (just previously accommodated and verged on a distant stimulus) the instant after a near stimulus has been turned on. (b) Equilibrium on the near stimulus. (c) Equilibrium on the near stimulus disturbed by sudden placement of a positive lens before each eye. (d) Equilibrium achieved in viewing the near stimulus through the positive lenses; 'accommodation disparity' is $\triangle a_1$, 'vergence disparity' is $\triangle v_1$.

is no initial 'resistance' of the g_v template to the movement because the vergence marble was resting at the bottom of an energy well where the slope was zero. At some point before accommodation reaches the bottom of the energy well in e_a, equilibrium is reached, as illustrated in Figure 8d.

The equilibrium position illustrated in Figure 8d has two interesting properties. First, the final equilibrium position may not be uniquely determined, because the marbles come to rest when they both lie on e_a and e_v slopes that are zero. This situation may occur for a significant range of possible values of a and v, with a correspondingly wide ambiguity of the 'equilibrium' values of a and v. Second, the equilibrium positions of a and v occur at the minima of e_a and e_v, and these minima are displaced from the corresponding minima of h_a and h_v. The displacement of the minima of e_v from those of h_v was previously noted to be the cause of 'vergence disparity.' A similar displacement of the minima of e_a from those of h_a is the cause of 'accommodation disparity.'

At the instant the +1 diopter of 'accommodation disparity' was introduced by the spectacles, the sum of the 'vergence disparity' plus 'accommodation disparity' was 1.0. As the overall system approaches the equilibrium position (Figure 8d), the total disparity is continuously reduced, but it can never quite reach zero; some 'vergence disparity' and some 'accommodation disparity' remain at the minima of h_a and h_v whenever $a \neq v$. The final equilibrium depends complexly on all of the relevant factors: the stimulus to accommodation, the stimulus to vergence, and the exact shape of the accommodation and vergence displacement functions.

Interaction of Fusion and Vergence

In the model, the effect of vergence on fusion is simple: vergence determines the horopter and hence the center of the g_u displacement function.[17] So far, however, the model does not contain a corresponding effect of fusion on vergence. The computation of image disparity, which is virtually identical for fusion and vergence, suggests a possible action of fusion on vergence.

The speed of vergence is restricted by the mechanical properties of the eyeball and its muscular apparatus. Fusion is not so restricted. Consequently, fusion of an object can occur before vergence upon it is complete—or, for that matter, before vergence has even begun. The model of vergence is based on $\partial e/\partial v$ (Equation 1), which is calculated in an infinitesimal neighborhood of v (Equation

---

[17] The effect of vergence on perceived depth is far more complex, and beyond the scope of this theory. See K. N. Ogle, Spatial localization through binocular vision, in H. Davson (ed.), *The Eye*, 4, 1962, 271–324; W. Richards and J. F. Miller, Convergence as a cue to depth, *Percept. & Psychophys.*, 5, 1969, 317–320.

6). The definition of e_v does not take any account of the information the fusion system may have about vergence possibilities up to ±2 deg. from v. Therefore, *a new vergence image-disparity energy function* (H_v) is proposed here, to incorporate fusion information.

In defining H_v, the h_v definition (Equation 6) is modified to give fused stimuli greater importance than unfused stimuli in determining vergence, by incorporating a fusion factor $[U(x, y, u)]$ into the defining integral:

$$H(v, U) = \iint_R [|l(x - v/2, y) - l'(x + v/2, y)| + U(x, y, v)]w(x, y)\, dx\, dy. \qquad [11]$$

The fusion factor is $U = 0$ when fusion has occurred at the plane $v = u$ in a neighborhood of x, y; and $U > 0$ when fusion has not occurred. In general, $U(x, y, v)$ is assumed to vary with distance from the plane of fusion in roughly the same way that $g_u(\Delta u)$ varies with Δu (Figure 6). The revised definition of H_v by Equation 11 changes nothing of fundamental importance in the treatment. It merely makes e_v a more accurate representation of vergence by incorporating an interaction of fusion with vergence.

Before concluding this section, it must be mentioned that horizontal vergence, and horizontal convergence in particular, can—to a much greater extent than vertical vergence, fusion, or accommodation—be controlled voluntarily with practice. Voluntary control is represented as a force, $\partial e_v/\partial v$; f_{vol} has been neglected in the theoretical treatment, but it must be considered, particularly in composing demonstrations that involve convergence.

Interaction of Accommodation and Fusion

The model contains an implicit image-term interaction of a and u; that is, failure to accommodate blurs an image and affects fusion of it. The model contains no provision for u to affect a, and there are no data to suggest that one is needed. Nor does the author know of any authentic three-term interactions (i.e., interactions not derivable from two-term interactions).

The Complete Physical Model: An Illustrative Example

Naturally occurring objects provide highly correlated stimuli to a, v, and u; they challenge neither the visual system nor the model.

However, images viewed in a stereoscope frequently not only challenge but defeat the visual system's attempts to cope with them. The model now undertakes to joust with one such stimulus.

Figure 9 illustrates a typical stereoscopic stimulus (*L* and *L′*) and the intended depth illusion (seeing a raised disk above a rectangle). The illusion of stereoscopic depth is difficult to achieve when the rectangle and disk are separated by about 1 diopter (≈3.5 deg.). Most observers readily can fuse that portion of the scene for which accommodation is appropriate—say, the rectangle—but they cannot at first fuse the disk. After some 'trying' (which consists of looking at various parts of the display), they suddenly are able to fuse the disk also. Once this fusion has occurred, they can look from the rectangle to the disk without disturbing fusion of either.

Figure 10 shows the model for the sequence of events just described. The complete model of Figure 10 differs from the model of Figure 8 only in that templates representing fusion have been added. The fusion templates are carried along with vergence, so that

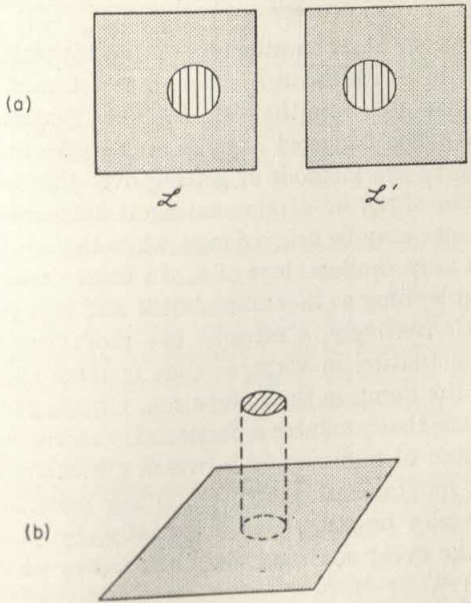


FIG. 9. A stimulus (a) to be viewed in a stereoscope; *L* and *L′* respectively are the images seen by the left and right eyes. (b) Perspective representation of percept induced by the stimulus in (a).

g_u always lies centered under g_v. The image-disparity functions $h_u(u, z)$ that add to g_u remain fixed, of course, and do not move. The two fusion templates illustrated represent two small areas of the retina, one in the central fovea, one in the periphery.

Figure 10a shows the starting position of the model when the picture is turned on. Immediately before, the observer has been looking at the frame of the stereoscope. Accommodation and vergence therefore are appropriate for the background rectangle. The disk, which falls on the fovea, is unfused and the eyes are not verged on it. The vergence image disparity (h_v) is indicated according to its original definition in Equation 6. The weighting function (w) assigns less weight to the peripheral stimulus (rectangle) than to the foveal stimulus (disk). The 'unweighting' of *peripheral* image disparity is confined to vergence and does not occur in the h_u fusion image-disparity function.

When confronted with unfused foveal images and fused peripheral images, the eyes usually turn to fixate the fused image. Figure 10b illustrates this situation, with the fovea pointed to one side of the rectangle. There now are excellent stimuli to accommodation, vergence, and foveal fusion; only peripheral fusion of the disk is unsatisfied.

Eventually, as the observer attempts to fuse his entire visual field, his eyes return to fixate the disk and he has returned to his initial state (Figure 10a). To fuse the disk, the eyes first must converge. Convergence may be triggered in numerous ways. In terms of the model, these ways are methods of getting over the 'hump' between the two minima of e_v, or of eliminating it temporarily. In some cases, convergence may be delayed even when there is no hump, if it is slowed by a very shallow slope of e_v. In these cases, the observer need only keep looking at the unfused disk and his eyes will slowly verge on it. Alternatively, a saccadic eye movement may cause a spontaneous fluctuation in vergence that is large enough to bring vergence over the hump in the e_v function. Or, the eyes may find a location to fixate that produces a particularly strong vergence stimulus (a minimum of minima of h_v), which eliminates the e_v hump. Voluntary or spontaneous vergence-hunting motions may occur. Vergence may also be stimulated artificially, by placing negative lenses before the eyes; or it may be simulated by placing prism(s) before the eye(s).

For whatever reason, when the vergence marble finally arrives at the edge of the disk's e_v energy well, it rolls into it. Fusion of the

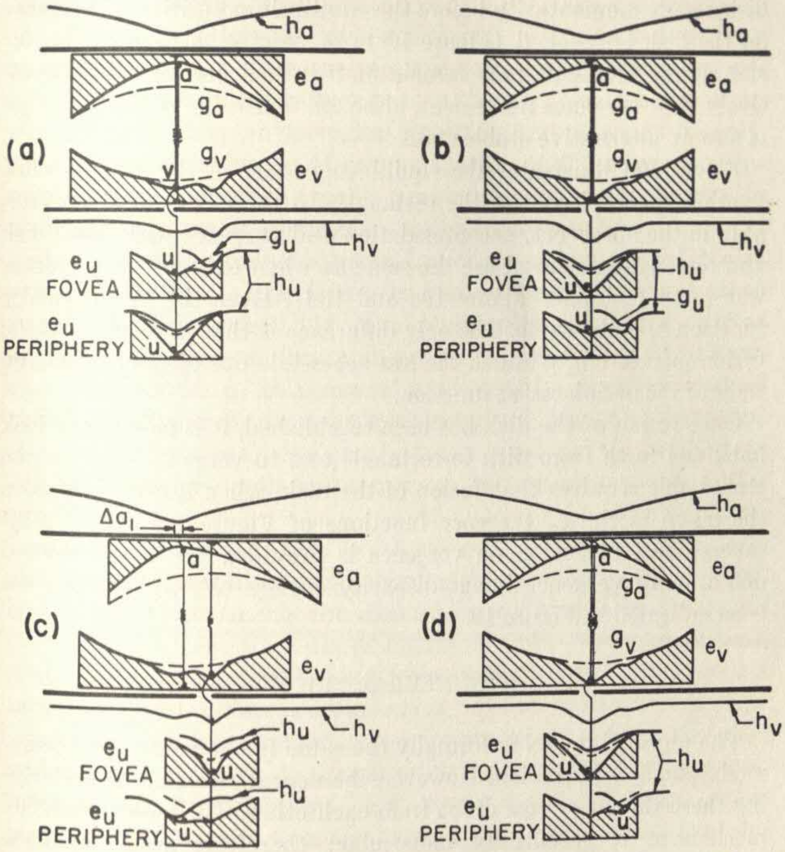


FIG. 10. Complete model of vergence, accommodation and fusion. The g_u fusion templates are attached to the v marble and move with it (as does g_a). The fusion image-disparity energy function $h_u(u, z)$ remains fixed. It adds to $g_u(u)$ to produce the e_u surface upon which the u marble rolls. A constant force (mechanism not indicated) propels the u marble downward. Each g_u template represents a different small area of retina. The minimum of h_u is drawn narrow for the disk, wide for the rectangle (see Figure 9). The h_v function is the sum of h_u functions, with the peripheral h_u reduced by $\frac{1}{2}$. (a) Initial viewing of the images of Figure 9a, fixation on disk: vergence and accommodation are appropriate for rectangle; accommodation is appropriate for disk also, but vergence is not and the disk is not fused (v and foveal u marbles are not in energy wells of the disk). (b) Fixation on rectangle: the disk, which still is unverged upon and unfused, now is in periphery. (c) Fixation on disk after vergence upon and fusion of disk: note that peripheral fusion of rectangle is maintained; note 'accommodation disparity' Δa_1, and negligible 'vergence disparity.' (d) Fixation on rectangle after (c): note that peripheral fusion of disk is maintained.

disk occurs momentarily before the equilibrium position of vergence on the disk is reached (Figure 10c). In verging to equilibrium, fusion of the rectangle may remain undisturbed because of the relatively slow vergence movement, although nonfusion of the rectangle is now an alternative stable state.

Figure 10d illustrates the equilibrium state of the system when fixation is transferred to the rectangle. In the model (and presumably in the observer), accommodation and vergence when both disk and rectangle are fused are the same as when only one or the other was fused (compare Figure 10a and 10d). Even the fusion energy function e_u is the same. The only difference is that in Figure 10d the fusion marble has attained the less accessible one of the two energy wells of the peripheral e_u function.

Once fusion of the disk has been established, it is possible to look back and forth from disk to rectangle and to verge on each. In the model, this requires that fusion of the disk, when it occurs, remove the hump in the e_v vergence functions of Figure 10a. This facilitating effect of fusion on vergence is provided for in the revised definition of vergence image-disparity energy (H_v), but it has not been indicated in Figure 10.

THE NEURAL MODEL

The physical model is formally the same for vergence, for accommodation, and for fusion. However, the neural mechanisms underlying these three systems differ from each other. The accommodation mechanism is essentially monocular; therefore, accommodation itself is of interest here mainly in terms of its interaction with vergence. (A neural mechanism of accommodation is proposed in Appendix A.) On the other hand, fusion and vergence are essentially binocular processes; both require the interaction of information from both eyes. The proposed neural process is basically the same for vergence and for fusion; the main difference being that fusion is computed separately in each region of the retina whereas vergence signals are summed over the whole retina. Thus, fusion serves as the kernel of the vergence system. In addition, fusion involves binocular rivalry and other interesting phenomena. The purpose of the neural model is to propose an analysis of the component processes of fusion; the extension to the processes of vergence is merely outlined.

The analysis is presented in terms of the functions performed by hypothetical neurons. The hypothetical neurons resemble real neurons, but whether they resemble neurons of the visual system is an unanswered question. The analysis is entirely functional, so any exact resemblance would be fortuitous. Although there are important new reports of electrophysiological recording from binocularly sensitive neurons in the brains of cats and monkeys,[18] the neurophysiological data presently available are too fragmentary to provide either confirmation or disproof of the proposed model.

The core of the proposed neural mechanism is a *neural binocular field* (*NBF*) that performs the kind of binocular comparisons described by Equation 7. The use of a binocular-comparison field as an explanatory concept for depth perception can be traced to Kepler;[19] neural fields of this general kind are also familiar in other contexts.[20] Psychophysically, the existence of binocular point-by-point comparison has been amply documented by the demonstration of binocular depth perception in the absence of monocular cues.[21]

As diagrammed in Figure 11, the neural binocular field (NBF) is an internal three-dimensional representation of the external world. The x, y plane represents the dimensions perpendicular to the observer's line of view; the z dimension represents depth. At each level in the NBF, signals originating from the left and right eyes are represented with different translations relative to each other. Signals originating from corresponding points of the retinas intersect in the middle of the field; signals originating from points displaced temporally in the two retinas (i.e., signals produced by an object nearer to the observer than his fixation point) intersect below the middle of the field; signals from a more distant object intersect at the top of the field.

---

[18] H. B. Barlow, C. Blakemore, and J. D. Pettigrew, The neural mechanism of binocular depth discrimination, *J. Physiol. (London)*, 193, 1967, 327–342; T. Nikara, P. O. Bishop, and J. D. Pettigrew, Analysis of retinal correspondence by studying receptive fields of binocular single units in cat striate cortex, *Exp. Brain Res.*, 6, 1968, 353–372; D. H. Hubel and T. N. Wiesel, Cells sensitive to binocular depth in area 18 of the macaque monkey cortex, *Nature*, 225, 1970, 41–42.

[19] See reviews in E. G. Boring, *The Physical Dimensions of Consciousness*, 1933 (reprinted 1963), and in L. Kaufman, On the operations underlying stereoscopic combination, SP-66-1, 1966, Sperry Rand Research Center (Sudbury, Mass.).

[20] W. S. Pitts and W. McCulloch, How we know universals, *Bull. math. Biophys.*, 9, 1947, 124–147.

[21] B. Julesz, Binocular depth perception of computer-generated patterns, *Bell Syst. tech. J.*, 39, 1960, 1125–1162.

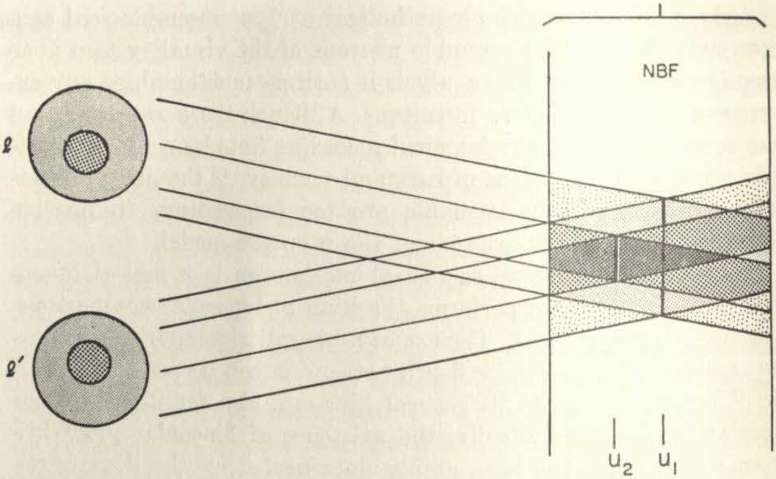


FIG. 11. Diagrammatic illustration of the Keplerian theory of stereopsis. A cross section of the neutral binocular field is indicated by NBF, at the right. Neural signals originate from images on the left and right retinas (*l* and *l′* respectively) and pass through the NBF along the straight lines drawn from *l* and *l′* to the NBF. When a region of the field receives inputs from both *l* and *l′*, these inputs interact. Regions of interaction of the center and outer disks respectively are indicated by heavy and by medium stippling. Vertical lines, u_1 and u_2, indicate levels of the NBF at which exact binocular correspondences occur between the signals from outer and inner disks respectively of *l* and *l′*. These two different levels represent different distances from the observer (depth). The figure is so drawn that the direction of signal flow can be reversed. In this case, levels of binocular correspondence u_1, u_2 in the NBF are interpreted as cross-sections of external stimuli. The signals (light) flow from right to left to produce the images *l* and *l′*. This method of representation demonstrates how, in Keplerian theory, the contents of the NBF are a simple reflection of the external world.

Up to this point, the model reiterates the general ideas of Koffka, Charnwood, Linksz, Julesz, Dodwell and Engle, Kaufman, and many others.[22] The model presented here embellishes on its predecessors in three respects. First, it contains two NBFs: a primary NBF for the fine-detail functions of binocular vision, and a secondary NBF for coarse-detail functions. The interaction of these two systems is crucial for many phenomena of binocular depth perception. Second, the model analyzes binocular fusion into elementary subfunctions, performed by hypothetical neurons. Third, it accounts for fusion and for rivalry within the same system and in the same neurons; it suggests how neurons that yield binocular fusion in some circumstances can yield binocular rivalry in others.

---

[22] K. Koffka, Some problems of space perception, in C. Murchison (ed.), *Psychologies of 1930*, 1930, 161–187; J. R. B. Charnwood, *Essay on Binocular Vision*, 1951; A. Linksz, *Vision, Physiology of the Eye*, Vol. 2, 1952; P. C. Dodwell and G. R. Engle, A theory of binocular fusion, *Nature*, 198, 1963, 39–40, 173–174; L. Kaufman, Some new stereoscopic phenomena and their implications for the theory of stereopsis, this JOURNAL, 78, 1965, 1–20.

Primary Neural Binocular Field

Inflow neurons. Figure 12 illustrates the four kinds of neurons proposed to constitute the primary NBF. The first of these are the inflow neurons (Figure 12a), which carry signals that originate from the left and right retinas. Their axons cross the NBF at opposite slants. Inflow signals are carried by their axon terminations to all depths in the field. The inflow neurons are assumed to con-

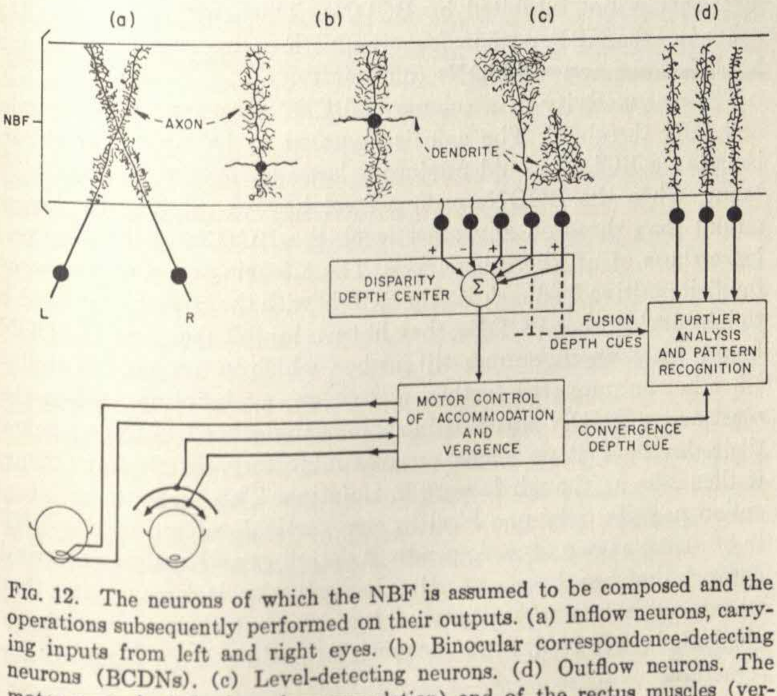


FIG. 12. The neurons of which the NBF is assumed to be composed and the operations subsequently performed on their outputs. (a) Inflow neurons, carrying inputs from left and right eyes. (b) Binocular correspondence-detecting neurons (BCDNs). (c) Level-detecting neurons. (d) Outflow neurons. The motor control of the lens (accommodation) and of the rectus muscles (vergence) is indicated at lower left.

nect to all other neurons in the NBF. Figure 12 is drawn so that the number of synapses (connections) between axons (outputs) of one neuron and dendrites (inputs) of another neuron in a given small area is proportional to the product of the density of axon terminations and dendrites within that area.

Binocular correspondence-detecting neurons (*BCDNs*). A BCDN-(Figure 12b) receives its input from only one level of the NBF and sends its output to all other levels. The function of the BCDNs is to limit outflow from the NBF to just one particular level when a binocular correspondence exists at that level. A BCDN receives inputs from inflow neurons, and it is silent unless its input indicates a correspondence between the left and right retinas. When its input does indicate correspondence, it produces a large output; this output acts to potentiate outflow at its own level and to inhibit BCDNs on all other levels. The precise nature of the inputs is considered below, in the section on rivalry. The transfer of signals from inflow neurons to outflow neurons is assumed to be possible only at the level that is not inhibited by BCDNs. Thus, within any small z column of the NBF, outflow from the NBF occurs only from the one level that has active BCDNs (monoactivity).

The monoactivity of a column of BCDNs depends on their gain and their threshold. The gain is assumed to be large; that is, in isolation a BCDN would produce a large output in relation to its input. When the BCDNs on one level have a significantly larger output than those on any other level, the BCDNs of this level silence those of all the other levels. The silencing occurs because of implicit positive feedback. As the BCDN with the largest input begins to inhibit the other BCDNs, they in turn inhibit the original BCDN less; it increases its output still further, which in turn acts to diminish other outputs still further, and so on, ad infinitum. When the most active BCDN succeeds in suppressing other BCDNs to below their threshold, it no longer receives inhibitory signals from them; it then acts as though it were in isolation. This monoactive interaction permits only one level in any vertical column of the NBF to be active at one time. Appendix B describes and analyzes a neural network that has these properties; in particular, it demonstrates that when neurons are appropriately interconnected, their normal physiological characteristics are adequate to produce the monoactive interaction.

If it is assumed that BCDNs in the middle level of the NBF have a slightly lower threshold (or equivalently, receive a slightly higher level of spontaneous activity) than other BCDNs, then when only one eye is stimulated, the middle level is the only active level. The bias in favor of the NBF's middle level corresponds to the g_u function of the physical model for fusion; recall that $g_u(\Delta u)$ expressed the tendency for fusion to occur at $u = 0$ (middle level) in preference to other depth planes.

In general, the computation performed by the BCDNs corresponds to the $l - l'$ comparison operation of the h_v and h_u image-disparity functions (Equation 6) of the physical model. Different levels of the NBF correspond to different values of Δv and Δu at which $l - l'$ is computed. The active level of the NBF corresponds to the plane of fusion of the physical model; the outflow from this level corresponds to the subjective binocular image.

Level-detecting neurons. The third kind of neuron detects the active level of the NBF, that is, the depth plane at which binocular correspondence occurs. The function of the level-detecting neuron could be combined with that of the BCDN, but I have chosen to designate it separately. Level-detecting neurons (Figure 12c) receive their input from BCDNs. The number of connections between a BCDN and a level-detecting neuron increases with the BCDN's distance from the center of the NBF, so that a binocular correspondence at an extreme level of the NBF produces a big input to the level-detector and hence a big output.

The outputs of level-detecting neurons are evaluated at a disparity-depth center. Level-detecting neurons that receive signals primarily from the bottom of the NBF feed excitatory signals to the disparity-depth center; neurons from the top of the NBF feed inhibitory signals. Thus the inputs to the disparity-depth center indicate the depth levels of binocular correspondences.

Information about the fusion depth of each small area of the NBF is passed through the disparity-depth center to subsequent stages for further processing. In addition to transmitting local depth information, the disparity-depth center sums the combined outputs of the level-detecting neurons over the NBF. This summed output signal (plus the equivalent output of the secondary NBF) is used to control vergence; it is the net contribution of retinal disparity to the control of vergence. In the physical model, the image component of the *force* that controls vergence is $\partial h_v/\partial v$ (the image's contribution to the slope of the e_v surface under the vergence marble); this force corresponds to the summed output of the disparity-depth center.

As defined, the disparity-depth center produces a signal inversely proportional to depth, that is, a convergence signal. As a rule, the nervous system provides functions in paired opposites; therefore, a more complete diagram would include an oppositely signed depth center to provide a divergence signal. The factor g_v of the physical model, which represents motoric and all influences on vergence other than disparity cues, is implicitly represented in Figure 12 by the box labeled Motor Control.

Outflow neurons. The outflow neurons (Figure 12d) continue the analysis of shape, texture, and color in the NBF. These neurons are organized perpendicularly to the NBF and they receive inputs equally from all depths. Their input is significantly simplified by the BCDNs, which limit the inflow-outflow transfer to a single level of the NBF. If it were not for BCDNs, the outflow neurons would receive the summed signal of various planes of comparison, a blur of useful and useless information. The outflow neurons carry no binocular depth information. At subsequent stages of analysis, however, binocular depth information (from the level-detecting neurons) is recombined with the shape, texture, and color information of the outflow neurons.

We may note here that the neurons which control accommodation are not represented in Figure 12. To provide the h_a defocus signal, any combination of contour-sensitive neurons (e.g., on-center, off-surround) whose output was added at an 'accommodation center' would suffice (see Appendix A).

'Binocular-rivalry neurons': BCDNs I–IV. When antagonistic stimuli are presented to the two retinas (e.g., l = horizontal, l' = vertical stripes), then in some areas of the visual field an observer sees the stimulus from one eye; in the remaining areas he sees the stimulus from the other. For many stimulus pairs, there are no areas of the field in which stimuli from both eyes are seen simultaneously. This 'either-or' interaction between stimuli to the left and right retinas is called binocular rivalry, a phenomenon so closely related to binocular depth perception that one can hardly discuss one without the other. Here, the features of the NBF are elaborated to account for these rivalry interactions.

Only one generic type of BCDN has so far been proposed, to detect binocular correspondences in the NBF. Here it is further proposed that the NBF has *four* generic types of BCDNs, organized as two pairs, and that reciprocal interactions between these pair-mates on the *same* level of the NBF account for binocular rivalry. (The interactions between BCDNs on different levels are as described above: an active level suppresses all other levels, and this interaction accounts for fusion and depth phenomena.)

To illustrate the four BCDN types, a nomenclature for describing neuronal inputs is helpful. Let F represent the function transmitted by an input, and let $\sim F$ represent the complementary function. For example, F may represent input to the BCDN from an on-center, off-surround receptive field as described by Kuffler and by Hubel;[23] $\sim F$ then represents an off-center, on-surround receptive field. Other examples of antagonistic receptive-field functions would be a vertical light-dark boundary (F) and a dark-light boundary ($\sim F$), a vertical bar (F) and a horizontal bar ($\sim F$), and so forth. Further, let L represent an input from the left eye and R an input from the right eye. Let + represent an excitatory input connection and − an inhibitory input connection. A BCDN is inactive when it receives no input signals or when it receives only inhibitory signals. When it receives a signal through an excitatory input, it becomes active. If both excitatory and inhibitory signals are received, the activity is less than that produced by only excitatory signals.

Table I lists the four kinds of BCDNs for a particular function (F). Types I and IV are F detectors; they are active when the eyes receive a stimulus satisfying the function (F)—in this example, a white spot on a dark surround. When only the left eye receives the

TABLE I

INPUTS OF THE FOUR TYPES OF BCDNs

Type	Eye	Message	Input
I	L	F	$(F, L, +)$ and $(\sim F, R, -)$
II	L	$\sim F$	$(\sim F, L, +)$ and $(F, R, -)$
III	R	$\sim F$	$(F, L, -)$ and $(\sim F, R, +)$
IV	R	F	$(\sim F, L, -)$ and $(F, R, +)$

Pair *A*, BCDNs I and III; Pair *B*, BCDNs II and IV.

---

[23] S. W. Kuffler, Discharge pattern and functional organization of the mammalian retina, *J. Neurophysiol.*, 16, 1953, 37–68; D. H. Hubel, Integrative processes in central visual pathways of the cat, *J. opt. Soc. Amer.*, 53, 1963, 58–66.

stimulus, Type I is active; when only the right eye receives the stimulus, Type IV is active. Types II and III are the corresponding $\sim F$ detectors.

All four types of BCDNs are assumed to be present in every small region of every level of the NBF. The function of the BCDNs is to detect binocular correspondences and then, by means of strong inhibition of BCDNs at all other levels, to restrict outflow from the NBF to the level at which the correspondence occurred. The four types of BCDNs are assumed to *enable specific* inflow-outflow connections: left-eye BCDNs (Types I and II) enable transfer of information from left-eye inflow neurons to outflow neurons, and right-eye BCDNs (Types III and IV) enable right-eye inflow-outflow transfer.

The four BCDN types are organized into *strong reciprocally inhibiting* pairs within any given level of the NBF. Two kinds of BCDN pairs are proposed here: A pairs, composed of BCDN Types I and III; and B pairs, composed of BCDN Types II and IV. The two partners in each pair receive identical inputs except that excitation and inhibition are reversed. Within a level the two members of a pair interact in basically the same way as two different levels interact: a member is fully active or silent, and only one member of the pair can be active at one time. The active member is usually the member with the larger input, although previously established activity may persist even when the inputs are changed. When neither member has an input that exceeds the threshold (ϵ), then both members of the pair are silent.

Finally, it is assumed that—in addition to their strong reciprocal interconnections with their partners in a pair—the four kinds of BCDNs have similar but much *weaker reciprocally inhibiting* interconnections with neighboring opposite-eye neurons on the same level. For example, BCDN I inhibits neighboring BCDN IIIs and IVs weakly, in addition to strongly inhibiting its BCDN-III partner. The effect of this interaction is facilitory, in the sense that as soon as one member of one pair has achieved dominance, it helps to silence ambivalent opposite-eye neighbors and thereby to facilitate the quick spatial spread of dominance of BCDNs of the same eye throughout that level of the NBF.

To illustrate the operation of BCDN pairs, we consider first the case where both eyes receive the same stimulus. Then, one BCDN of each pair receives an excitatory input, and the other member receives an inhibitory input. For example, when the stimulus on both retinas satisfies F (when it is, say, a bright spot on a dark background), the left-eye input of BCDN I (which has an excitatory connection) is excited, and the right-eye input of BCDN I (which has an inhibitory connection) is inhibited and carries no signals. Therefore BCDN I is excited. The excitatory input of BCDN III is silent (it requires a dark spot), and the inhibitory input is active; therefore, BCDN I is active; BCDN III is inhibited. In Pair A then, BCDN I dominates BCDN III. Similarly, in Pair B, BCDN IV dominates BCDN II. The active BCDN I's enable left-eye inflow-outflow connections in the neural field, and the active BCDN IVs enable right-eye inflow-outflow connections; all inflow-outflow connections are enabled. And when the stimulus on both retinas satisfies $\sim F$ (when it is a dark spot on a light surround), BCDN II dominates I and BCDN III dominates BCDN IV; again, all inflow-outflow connections are enabled.

When the stimulus to one retina satisfies F and the stimulus to the other retina satisfies $\sim F$, however, then only one set of inflow-outflow connections is enabled. For example, suppose that the left eye receives a bright spot and the right eye receives a dark spot. BCDN I receives excitatory signals from the left eye and inhibitory signals from the right eye. BCDN III has the opposite inputs. Which of BCDN I and III dominates Pair A depends on the relative strength of the two stimuli, on past stimulation, and on bias factors. However, only one member can be active at one time. If BCDN I is the active member of Pair A, then the left eye's inflow-outflow connections are enabled; if BCDN III is active, the right eye's inflow-outflow connections are enabled. Normally, Pair B would enable the connections of the eye that was not enabled by Pair A. In the case of antagonistic stimuli ($L = F$; $L' = \sim F$), however, both members of Pair B (both BCDN II and BCDN IV) fail to receive excitatory input signals. Therefore neither member of Pair B is active, and only one eye's inflow-outflow connections are enabled—those of the eye chosen by Pair A.

Secondary Neural Binocular Field

The secondary NBF is assumed to be functionally similar to, and composed of the same types of neurons as, the primary NBF, the main difference being that the sizes of the receptive fields of its input neurons are about 10 to 20 times greater. The depth range of the primary NBF is estimated to be about $\pm\frac{1}{8}$ deg., that of the secondary NBF about ±2 deg. These depth ranges are very approximate characterizations of central vision; in both NBFs the depth range increases rapidly with peripheral angle.[24] The depth signals from the primary and secondary NBFs add in proportion to the magnitude of the depth; thus the primary NBF signals are merely a perturbation on the larger depth signals from the secondary NBF.

The fine-detail output of the primary NBF contains information about stimulus shape and texture, information which is used in subsequent decoding of these stimulus properties, such as in pattern recognition. The outflow of the primary NBF ultimately determines the contents of perception—*what* we see. The coarse-detail output of the secondary NBF—except for the depth signal—does not serve pattern recognition so directly. Metaphorically, we may say the secondary NBF is concerned with the metalanguage of perception; it directs recognition processes to particular regions of the visual field, but it is the outflow of the primary NBF that is processed. Because the information dealt with by the secondary NBF is so coarse, the secondary NBF computes binocular correspondence between stimuli that would be treated as different by the primary NBF. For example, at reading distance, the primary NBF clearly discriminates between the letters A and B; the secondary NBF treats them as identical 'blobs.'

Predictions of the Neural Model

There are two kinds of outputs from the primary and secondary NBFs that can be related to behavioral data: the *depth signal*, which derives from both NBFs, and the *pattern outflow*, which derives from the primary NBF. Below, the depth signals for fusion and vergence derived from the neural model are compared with those derived from the physical model; pattern outflow is considered next; and the combinations of depth signals and pattern outflow are considered last.

Depth signals: Fusion. The active level of an NBF corresponds to the level (Δu) at which fusion occurs in the physical model; a signal indicating this level is transmitted by the level-detecting neurons (Figure 12c). The computation of binocular correspondence by the BCDNs corresponds to h_v (and h_u) in Equation 6 of the physical model. The layer of BCDNs with greatest binocular correspondence corresponds to the fusion value of the h_u function in the physical model, and the spontaneous input to the middle layer of the NBF corresponds to the g_u bias function. The addition of g_u and h_u to form e_u in the physical model is represented in the neural model by the combination of the spontaneous and stimulus-produced inputs to the BCDNs; the particular energy well of e_u into which the marble representing fusion rolls depends on past stimuli and responses, just as does the particular active layer of the NBF.

---

[24] Ogle (n. 17).

The multistability of fusion in stabilized vision is the work of the BCDNs. Activity at any level of an NBF—once it is established—inhibits activity at any other level (monoactivity). The normal mechanism for establishing activity at a new level is a sudden change in the entire input, such as that caused by a saccadic eye movement or a blink, which in effect resets the NBF to a neutral starting position and allows the level which contains the strongest stimulus to establish itself.

In stabilized vision, activity can be established at a level (u) far from the center of the primary NBF by slowly shifting a fused L and L' laterally. Sooner or later, the level of maximal correspondence ($\triangle u$) will contain a shallower energy well than that produced by the bias in favor of the center level, but fusion at Δu continues to predominate until a sudden change causes a reevaluation of stimulation at all levels (see Figure 10d).

The physical model for fusion assumes that the image disparity of all possible fusion depths (u) is computed. The marble representing fusion rests in a particular energy well because of its past history. The neural suppression of one level in the NBF by another is equivalent to the suppression of certain fusion computations, that is, to the elimination of all but one of the energy wells—the one in which the marble rests. This suppression does not grossly disturb the isomorphism between the neural model and the physical model because the marble representing fusion can rest in only one energy well at one time, and therefore the presence or absence of the others in the neural model is not of first-order importance.

Depth signals: Vergence. The signals for controlling vergence are the summed depth signals from the primary and secondary NBFs, weighted according to the signals' retinal location, with foveal values having the greatest weight. The summing and weighting is carried out at the disparity-depth center (Figure 12). This signal (h_v) is combined with the motor-bias component (g_v) to give the net vergence-control signal (e_v).

The multistability phenomena illustrated in Figure 4 (that the eyes may fail to verge on a stimulus when it is presented suddenly, although they can verge on it when it is produced by slowly transforming a previously verged stimulus) result from the combination of motor-bias forces (g_v) with image-disparity forces (h_v). However, complete failure of vergence (Figure 5) results when the disparity of the stimuli to the left and right retinas exceeds that for which binocular correspondence is computed in the NBFs. This and other interactions of fusion with vergence were subsumed in the physical model by Equation 12. Vergence failure usually does not occur outside the laboratory because the link between accommodation and vergence provides a vergence stimulus for naturally occurring objects.

Pattern outflow: Binocular rivalry. The neural model predicts that when stimuli to the two retinas are in correspondence, the outflow from the primary NBF contains the total pattern message from both eyes. When the retinal stimuli are not in correspondence, the outflow from the primary NBF contains messages only from one eye; the chosen eye depends on the stimulus configuration (past and present), on ocular dominance, and on accident. For a given pair of rivalrous stimuli, the condition of one eye being dominant is a stable condition, which, in the model, would persist indefinitely until a perturbation such as a blink or an eye movement changed the inputs.[25] Insofar as activity tends to fatigue BCDNs, fatigue as well as perturbations of input would contribute to alterations of dominance.

By assuming that the function (F) used by BCDNs in the primary NBF is on-center and off-surround, many additional facts of binocular rivalry are accounted for. For example, when, in a given retinal neighborhood, one stimulus contains contours and the other does not, the stimulus with contours dominates. This observation is predicted by the model because F is a contour-sensitive function; the contour input to the BCDNs enables inflow-outflow transfer in its eye, whereas the noncontour input to the other eye does not enable transfer.

---

[25] N. J. M. Levelt, Note on the distribution of dominance times in binocular rivalry, *Brit. J. Psychol.*, 58, 1967, 143–145.

The model further predicts that a contour stimulus (say, a ring) presented to one eye enables inflow-outflow from the same eye not only in the neighborhood of the contour but also for the whole inside of the ring. It predicts this even for rings whose radius exceeds the range of lateral BCDN-pair interactions—see the paragraph on weak BCDN interactions—because lateral spread is a self-propagating phenomenon when no resistance is encountered (i.e., a rivalrous visual stimulus). Levelt summarizes many observations to prove that common contours enable binocular summation in the area within the contours.[26] For example, when a black disk on a white surround is presented to the left eye and a uniform field, white or dark, to the right eye, the disk dominates and is seen unaltered. But when a disk is presented to the left eye and a ring to the right eye, a disk is seen and its observed brightness is intermediate between that of the presented disk and that of the field within the ring. In the neural model, inflow-outflow is enabled within both the disk and the ring, so that the net outflow is the combination of both inflows, dark and light. The model does not specify how subsequent stages respond when they receive messages from the primary NBF that an area is both light and dark, but gray—the average—is a reasonable solution.

The model predicts in some detail what is seen in such complex stimuli as horizontal stripes to one eye and vertical stripes to the other. There are two areas at the intersection of the edges of the stripes where the BCDN pairs receive conflicting inputs from each eye (see Figure 13). Once one retinal stimulus (*l*), say, the vertical stripes, has dominated the other in one boundary conflict, lateral interaction tends to cause dominance of that stimulus to spread throughout the whole stripe and into any adjacent stripes. If the stripes are wide enough, however, the contours of horizontal stripes (*l′*) from the other eye can establish little 'islands' of dominance within the vertical stripes (Figure 13). Brightness signals from the two eyes are combined interior to these islands because both BCDN pairs are active. Where *l* and *l′* present conflicting contours to the two eyes, only one member of one BCDN pair is active. Figure 13 indicates these areas of rivalry.

---

[26] N. J. M. Levelt, *On Binocular Rivalry,* 1965, Institute for Perception RVO-TNO (Soesterberg, The Netherlands).

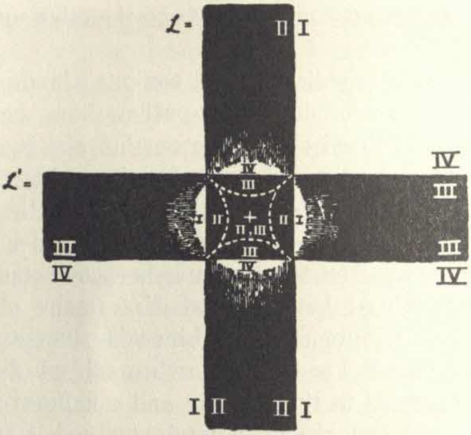


FIG. 13. Diagrammatic representation of the visual appearance of binocular rivalry. The stimulus to the left eye, *L*, is a vertical black bar on a white surround; the stimulus to the right eye, *L'*, is a horizontal black bar. The Roman numerals indicate the type of BCDNs stimulated in each area of the view. In the center and in the background, a BCDN member of each of the *A* and *B* pairs is stimulated; these areas are not seen in rivalry. In other areas, members of only one pair are stimulated; these are areas of rivalry. (After Helmholtz [n. 3], 496.)

The binocular rivalry discussed in the paragraphs above is within a depth plane. Rivalry *between* depth planes is another matter. Consider two objects, one behind the other, yet so placed that both objects are seen by each eye. The reader can arrange such a demonstration for himself with two pins. Stick the pins into a pencil about an inch apart, parallel to each other and perpendicular to the pencil. Hold the pencil against an illuminated surface at reading distance so that one pin is directly behind the other, as in Figure 14. As long as one pin is directly behind the other (pencil pointing between the eyes), both pins cannot be fused simultaneously; a double image is very obvious. When the pencil is rotated so that it points outside of the eyes, the angular distance between the depth planes of the pins is not changed much, but the pins can now be fused. This demonstration illustrates that even when both objects produce distinct retinal images in each eye (so there is no loss of information), an object cannot be seen as being behind another object at the same point of the visual field.

In the model, the inability to fuse two objects when one is directly behind the other results from the monoactivity of a column

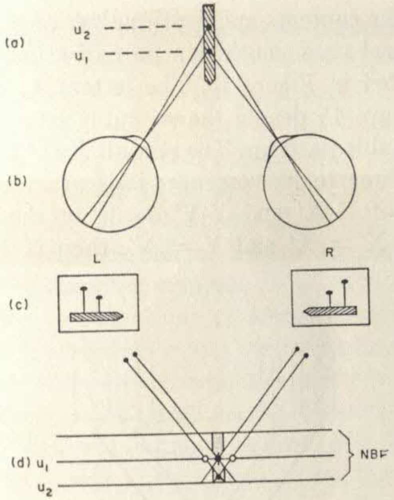


FIG. 14. A display in which two objects, one directly behind the other, nevertheless produce distinct images on each retina. (a) Top view of two pins in a pencil; u_1, u_2 indicate the depth planes. (b) Retinal projections of the stimuli. (c) Views of the stimulus as seen by each eye (these views are mirror images of the retinal projections). (d) Projection of the stimuli (c) onto the NBF: fusion of the near pin on level u_1 shown by an filled circle; the unfilled circles indicate heterogeneous fusion of a projection of the near pin with a projection of the far pin; shading indicates the area within which potential fusions are suppressed because of fusion at u_1; the filled circle at u_2 (in shaded area) indicates the point of binocular correspondence of images of the far pin, a point at which fusion is suppressed.

of BCDNs (i.e., outflow is restricted to only one level of the NBF). This rivalry between different levels of an NBF is fundamentally different from binocular rivalry between neurons at the same level of the NBF, where the within-pair interactions of BCDNs come into play.

Depth and outflow: Binocular combinations. A binocular view contains the outflow of the primary NBF (the contents) plus the combined fusion-depth values assigned to each x, y location by both NBFs. The binocular view is cyclopean because the two eyes combine to give a single outflow at each x, y location of the NBF; the contents of a binocular view therefore are basically the same as those of a monocular view. The important difference between monocular and binocular vision is the possibility of nonzero fusion-depth values in binocular vision.

To determine the contents and depth values of the cyclopean view in the most general case, consider a pair of stimuli (*L* and *L′*) of the type illustrated in Figure 15. The letters A, B, X, Y, and A′, B′, X′, Y′ in Figure 15 denote the stimulus letters A, B, X, Y, or any other identifiable patterns. The stimuli X-X′ and Y-Y′ function as a reference frame to fix vergence; it is assumed that binocular correspondences of X-X′ and Y-Y′ occur in the middle layer of the NBF. When X = X′ and Y = Y′, then X-X′ and Y-Y′ are

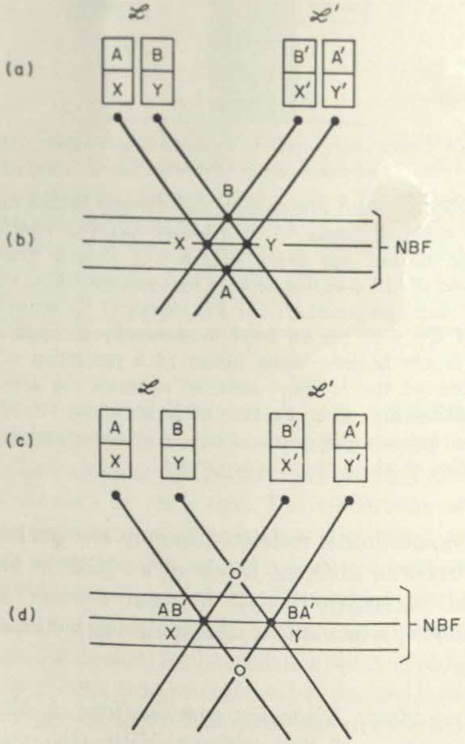


FIG. 15. Stimuli (a and c) to demonstrate binocular rivalry and/or depth; *L* and *L′* are the images to the left and right eyes respectively. The representations of the stimuli in the primary NBF are shown in (b) and in (d). (b) Monoactivity within each column of the NBF precludes seeing both A and B; only one is seen and in depth relative to X and Y. (d) A or B′ ($A \cap B'$) is seen together with $A' \cap B$, according to the outcome of rivalry in the central layer of the NBF. When binocular correspondence is detected in the secondary NBF (indicated by open circles), then $A \cap B'$ and/or $A' \cap B$ is seen in depth relative to X and Y.

each seen as a single stimulus at a depth determined by vergence and incidental cues. Which of A or B is seen? What is its depth and lateral position relative to that of X and Y? The answers depend on the nature of A and B, and on the lateral separation between them. The answers to these questions are completely general, because for any given region of any binocular stimulus, there is a choice of A, B, X, Y and of A′, B′, X′, Y′ that makes the reduced stimulus configuration of Figure 15 prototypical of any given region of the complex stimulus.

—*Binocular combinations of identical stimuli.* In the first group of cases, B is assumed to be a null stimulus (i.e., equal to the background). This eliminates complications due to binocular rivalry and means that the specific questions and answers involve binocular combinations of identical stimuli (A and A′). When the lateral separation between X and Y is very small (less than $\frac{1}{3}$ deg. in central vision), the binocular correspondence between A and A′ falls within the primary NBF (Figure 15b). In this case, a single stimulus (A-A′) is seen, above and behind X and Y, laterally between them. Not only does A-A′ appear laterally between X and Y, but, by varying the relative intensity of the stimulus to the left and to the right eye, the perceived lateral position can be made vary continously between its two extremes: with A laterally directly above X, as when only the left eye's stimulus of Figure 15a is presented; or A′ directly above Y′, as when only the right eye's stimulus is presented. This displacement effect is predicted by the physical model of fusion as follows.

Let u_1, u_2, u_3, respectively be the values of u at the minima of $g(u)$, $h(u)$, and $e(u)$. Whenever $u_2 \neq u_1$, then $u_1 \leq u_3 < u_2$ or $u_1 \geq u_3 > u_2$; u_3 is displaced from u_2 toward u_1 by the amount $\triangle u = u_3 - u_2$. The shift $\triangle u$ depends on the shallowness of the minimum of $h(u)$; the shallower the minimum, the greater $\triangle u$. The deepest minima of $h(u)$ occur when l and l' are of equal intensity; the minima become shallower when l is unequal to l' in intensity, and they vanish entirely when either l or l' vanishes. Changing the value of $\triangle u$ corresponds to sliding the point of fusion along one of the l-l' projections in the NBF, and hence to a lateral shift as well as a shift in depth. In fact, this lateral mobility of a fused object is the best empirical criterion of fusion in the primary NBF.[27]

---

[27] The question of whether fusion can cause a lateral displacement of the fused object has received conflicting answers in the literature. Large displacements invariably are the result of uncontrolled vergence and suppres-

Consider now what happens when we separate X and Y. When the lateral separation between A and A′ exceeds the $\frac{1}{3}$ deg. half-depth of the primary NBF, binocular correspondence between A and A′ cannot occur within the primary NBF. In this case, the middle layer of the primary NBF dominates; A and A′ are represented there separately (Figure 15d). Unless one eye is strongly dominant, A is seen above X-X′ and A′ is seen above Y-Y′. However, if A and A′ are large relative to the lateral displacement between them, they overlap and binocular rivalry may cause one to partially or completely suppress the other. For this reason, singleness of vision is not a sufficient criterion for binocular correspondence in the primary NBF.

When the lateral separation of X and Y is less than the 2 deg. half-depth of the secondary NBF, the apparent depth of A is determined by the depth signal from the secondary NBF. Suppose both A and A′ are seen, one above each of X and Y. Whether or not A or A′ or both appear in depth behind X-Y depends on their lateral separation, on the past history of stimulation, and on their size and texture. With large separations, two stable states (depth, no depth) are possible in the secondary NBF, but the depth sign of the A-A′ fusion in the secondary NBF may not always be attached to both the A and A′ representations in the primary NBF. Therefore, the same stimuli (A-A′) may appear behind X-Y on some occasions and in the plane of X-Y on other occasions. In brief, then, one or both images of A are seen because of their separate representations and possible rivalry in the primary NBF; neither, one, or both of the images appear in depth depending on whether binocular correspondence was detected in the secondary NBF. (These relations between the objects of Figure 15 are summarized in Table II, in the middle cell of the top row).

The last cell of the top row of Table II indicates the visual contents when the separation between X and Y exceeds the limit at which correspondences can be calculated in the secondary NBF. Then, in both NBFs, the only active layer is the middle one and no depth difference is seen between A and X-Y. Because A is represented twice in the middle layer, it is seen doubly. In all rows of

---

sion. This is demonstrated by briefly intensifying a steady *l* and *l′* alternately, and noting the reappearance of the suppressed elements. For background, see H. Werner, Dynamics in binocular depth perception, *Psychol. Monogr.*, 49, 1937, No. 2 (Whole No. 218); L. Kaufman, Suppression and fusion in viewing complex stereograms, this JOURNAL, 77, 1964, 193–205.

Table II depth disappears when the disparity exceeds the range of the NBFs, and in all cases it reappears (if it appeared in any cell of the row) when vergence movements are permitted.

Because the receptive fields of cells composing the secondary NBF are assumed to be very large, partial correspondences may be detected at disparities far exceeding the range of the secondary NBF. For example, suppose the extreme layers of the secondary NBF compute correspondences at disparities of 2 deg. If the diameter of receptive fields also were 2 deg., then stimuli separated by 4 deg. would still stimulate the edges of the receptive fields at the extreme layers, and a partial correspondence would be detected there. Brief flashes of simple stimuli separated by over 4 deg. do, in fact, yield appropriate localizations in depth (i.e., far or near) and they induce vergence movements in the appropriate directions,[28] but these results do not directly indicate the range of the secondary NBF.

—*Binocular combinations of rivalrous stimuli.* In this second group of cases, B is assumed to be a nonzero stimulus, different from A. This means that the question of what is seen involves binocular combinations of rivalrous stimuli (A and B—assuming that they are sufficiently different to 'rival' each other, a condition which it is obviously impossible to fulfill for very small X-Y separations). The results with rivalrous stimuli are basically the same as for stimuli that contain only A, except that wherever A was seen, only one of A or B is seen, depending on the local outcome of the A-B rivalry. These results are summarized in the second row of Table II.

When the difference between A and B is too small to be resolved by the secondary NBF (e.g., two different printed letters are interpreted as similar 'blobs'), then there are three possible levels of binocular correspondence in the secondary NBF: A-A′; A-B′, B-A′; and B-B′. Of these, the A-B′, B-A′ correspondence is most likely to dominate because it is on the middle level of the NBF. When the middle level dominates, relative depth between X-Y and A or B is lost. The last two cells of the second row of Table II summarize these results.

When A and B are so similar that they are taken to be identical

---

[28] D. E. Mitchell, Qualitative depth localization with diplopic images of dissimilar shape, *Vis. Res.*, 9, 1969, 991–994; G. Westheimer and D. E. Mitchell, The sensory stimulus for disjunctive eye movements, *Vis. Res.*, 9, 1969, 749–755; D. E. Mitchell, Properties of stimuli eliciting vergence eye movements and stereopsis, *Vis. Res.*, 10, 1970, 145–162.

TABLE II

CONTENTS OF THE CYCLOPEAN BINOCULAR VIEW FOR THE STIMULI SHOWN IN FIGURE 15

Stimulus	Disparity range (central vision)		
	Primary NBF ($\pm$ 1/8 deg.)	Secondary NBF ($\pm$ 2 deg.)	Beyond NBFs
$A = A', \ B = B' = 0$	$(d, AA', X/Y)$	$(nd \cap d, A, X) + (nd \cap d, A', Y)$	(nd, A, X) $+ (nd, A', Y)$
$(A = A') \neq (B = B') \neq 0$	$(d, AA' \cap BB', X/Y)$	$[(nd, A \cap B', X) + (nd, A' \cap B, Y)]$ $\cap [(d, A, X) + (d, A', Y)]$ $\cap [(d, B', X) + (d, B, Y)]$	$(nd, A \cap B', X)$ $+ (nd, A' \cap B, Y)$
$(A = A') \approx (B = B') \neq 0$	(nd, AB', X) $+ (nd, A' B, Y)$‡	$(nd, A \cap B', X) + (nd, A' \cap B, Y)$†	$(nd, A \cap B', X)$ $+ (nd, A' \cap B, Y)$†
$A = B = B', \ A' = 0$ (Panum)	(nd, AB', X) $+ (d, BB', X/Y)$	$(nd, AB', X) + (nd \cap d, B, Y)$	(nd, AB', X) $+ (nd, B, Y)$

Note: Each row represents a different angular separation (disparity) between X and Y. Each column represents a different stimulus content, A or B. The subjective visual contents are indicated by terms (extending from a left to a right parenthesis). Each term is composed of three elements. The first element represents depth: *d* indicates that the content (A or B) is seen in a different depth plane than X-Y: *nd* indicates no depth (i.e., the same plane). The second element represents content: $\cap$ indicates that either A or B is seen but not both; AB indicates that a concatenation of A and B is seen. The third element indicates location (lateral position) of the content: X/Y indicates that it is seen between X and Y; X indicates that it is seen directly above X; Y indicates that it is seen above Y; $X \cap Y$ indicates that it is seen above X or above Y but not both. To find the contents of a particular cyclopean view, sum the applicable terms. The table is not exhaustive, but all indicated combinations of elements are possible, though not necessarily for the same choice of A and B.

† Difference between A and B resolved by primary NBF, unresolved by secondary NBF.

‡ Difference between A and B unresolved by primary or secondary NBF.

not only by the secondary NBF but also by the primary NBF, then A fuses with B′ and B with A′ on the middle level of the primary NBF and no relative depth is seen. The first cell of the third row in Table II represents this situation. A-B′ fusion without depth is not the only stable state; but it takes some ingenuity to achieve the others. For example, if the intensity of B-B′ is initially set to zero, then fusion occurs on A-A′. By slowly intensifying B-B′, a second state (A-A′ fusion with depth) can be maintained, at least temporarily.

Because of its historical interest, Panum's limiting case[29] is described in the last row of Table II. In Panum's limiting case, the stimuli, A, B, and B′, are identical lines (e.g., the letter I), and A′ is null. The AB′ line is always fused in the X-Y plane. Whether the B-B′ binocular correspondence occurs in the primary or secondary NBF depends on the X-Y disparity, and it determines the lateral shift (if any) of B; in either case, it is seen behind X-Y.

The previously untested predictions of Table II have been verified experimentally by the author, using stimuli of the kind introduced by Kaufman following Julesz's method.[30] The demonstrations and the predictions have much in common with those of Verhoeff, Asher, and Kaufman.[31] It was Kaufman who first obtained and correctly interpreted evidence for a coarse-detail function (a 'blob' detector) in binocular correspondence detection.[32] Some of Kaufman's results are considered below; the neural model accounts for all the details.

Binocular combinations: The Kaufman figures. Figure 16ab illustrates two stereoscopic image pairs (L_1, L_1' and L_2, L_2') taken from Kaufman (n. 22). Only Figure 16a yields a stable stereoscopic depth illusion. Figure 16cd illustrates the representation of these stimuli in the secondary NBF. (Because of the large separation between characters, binocular correspondences at nonzero levels occur only in the secondary NBF.) Partial correspondences (any character of L with any other character of L') are indicated by open circles; exact correspondences (a character with its exact match) are indicated by filled circles. With the stimuli of Figure 16a, exact binocular correspondences occur only on one level of the NBF. When a BCDN becomes active on that level, activity quickly spreads to the limits of the region of exact correspondence, and beyond, suppressing the other levels.

---

[29] K. N. Ogle, Special topics in binocular spatial localization, in H. Davson (ed.), *The Eye*, 4, 1962, 349–407.

[30] Kaufman (n. 27); Julesz (n. 21).

[31] F. H. Verhoeff, A new theory of binocular vision, *Arch. Ophthal.*, 13, 1935, 151–175; H. Asher, Suppression theory of binocular vision, *Brit. J. Ophthal.*, 37, 1953, 37–49; Kaufman (n. 22).

[32] L. Kaufman, On the nature of binocular disparity, this JOURNAL, 77, 1964, 393–402; Kaufman (n. 22).

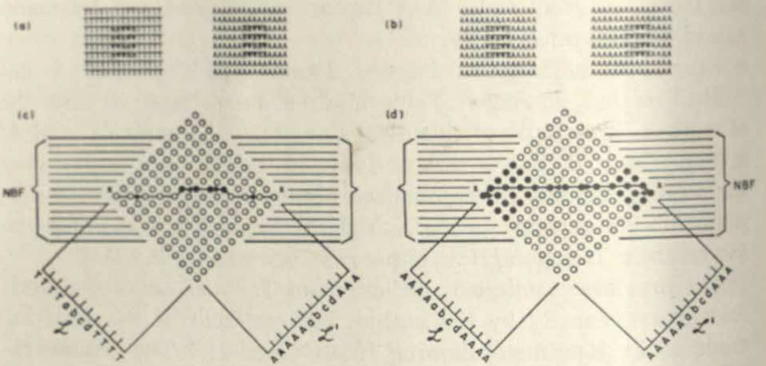


FIG. 16. Two stereogram pairs (a and b) from Kaufman (n. 22) and their representations (c and d) in the secondary NBF. Filled circles indicate exact correspondences, unfilled circles indicate coarse correspondences (i.e., of different letters); Xs indicate projection onto middle layer of outermost letters (for these, there is no binocular correspondence in the middle layer). Arrows indicate the direction of spread of activity within a layer; the lines through the arrows indicate the final, stable depth perception. For simplicity, only 12 of the 15 columns of the stereograms are indicated in (c) and (d).

Partial correspondence is possible at many levels throughout the secondary NBF. However, one BCDN that is always activated by a partial correspondence is the outermost one, because only this BCDN suffers no competition from correspondences at other levels. Since the uncontested, outermost partial correspondence is one level removed from the level of exact correspondence, it initiates a counterspread of activity throughout its level. The activity initiated by the partial correspondence at the edge spreads inward from the edge while the activity initiated by exact correspondences spreads outward from the center; ultimately they meet. The exact correspondence is a stronger input to the BCDNs than the partial one, so that wherever a conflict arises, the exact correspondence predominates. Therefore, in Figure 16c, the activity initiated by outer correspondences is excluded from the center rectangle. There are two stable states: the one illustrated in Figure 16c (which predominates) and the initial state (illustrated in Figure 16d).

For the stimuli of Figure 16b, activity initiated by an exact correspondence in the center of the stimulus spreads almost to the very edge (Figure 16d). Thus in Figure 16a, the center rectangle is seen in depth against its entire surround; in Figure 16b, only the end columns are seen in depth. By careful stereoscopic observation of Figure 16b the reader may be able to observe the temporal spread of activity from the center outwards: he also may observe briefly various semistable depth configurations.

The foregoing explanation accounts for the presence or absence of relative depth in Kaufman's demonstration. The content of what is seen (i.e., which letter) is determined in the primary NBF. Only the center level is active there because the disparity between adjacent letters exceeds its limits. Since the eyes verge on the inner letters, these fall on the center level. To note which letters are projected onto the center level of the primary NBF, the reader may refer to the center level of the secondary NBF (Figure 16c, 16d); there is no need for a separate diagram. For the stereogram of Figure 16a, which letter is seen depends on the A-T rivalry ($A \cap T$) at the center level of the primary NBF. The end columns are an exception because only the stimulus from one eye is represented at the center level of the NBF (crosses, Figure 16c). Thus the left-end column is A, and the right-end column is T.

In the stereogram of Figure 16b, the character A is seen at all its projections onto the center level. As for lateral position, all letters line up in vertical rows; the letters of the end columns (which can be seen in depth) do not appear to be shifted horizontally. (Lateral position is determined in the primary NBF, i.e., by the projections onto the center level in these illustrations.) These phenomena are subsumed in Table II in the second column (secondary NBF), the second ($A \neq B$) and third ($A \approx B$) rows. That is, A is partially confused with B in the secondary NBF and therefore the phenomena characteristic of both $A \neq B$ and $A \approx B$ occur. Most of Kaufman's other interesting demonstrations fall in the same cells of Table II.

Binocular combinations: The Hochberg stereograms. Hochberg published several interesting stereograms,[33] all of which use rivalrous stimuli with this property: that whenever binocular rivalry causes the image in one eye to dominate completely, the illusion of depth is lost. Since Kaufman's stereograms are counterexamples to this rule (only A or T is seen, yet depth persists; Figure 16a), we are faced with specifying the critical difference between the two kinds of stereograms. The main difference seems to be that the rivalry-inducing features of Hochberg's demonstrations are quite coarse, so that rivalry occurs in the secondary NBF as well as in the primary NBF.

---

[33] J. Hochberg, Depth perception loss with local monocular suppression: A problem in the explanation of stereopsis, *Science*, 145, 1964, 1334–1336;

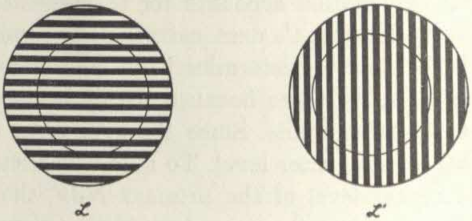


FIG. 17. Stereogram from Hochberg (third reference, n. 33) after Kaufman and Pitblado. The inner circles are seen in depth except when binocular rivalry of one pair member over the other is complete.

For example, to produce the stereogram of Figure 17, a horizontal grid is added to the left member and a vertical grid is added to the right member of a stimulus pair that would normally produce a stable illusion of depth. The relative coarseness of the grid feature explains the occasional loss of depth, because large conflicting features cause failure to compute the image-disparity minima in the secondary NBF.

Feature coarseness does not explain the correlation between suppression of one eye's input in the primary NBF and depth loss (suppression of an input in the secondary NBF). Significantly, however, the best operation for producing the correlated effects of suppression and depth loss is reducing the stimulus intensity to one eye.[34] Filtering the light input obviously affects both NBFs, which suggests that peripheral causes—such as vergence changes or other

---

J. Hochberg, On the conditions of stereopsis-suppression, *Science*, 146, 1964, 800; J. Hochberg, One view through two eyes: A theory of binocular combination and some supportive experiments, paper read at the meeting of the Psychonometric Society, October, 1964.

[34] Kaufman and Pitblado, cited by Hochberg (see the second reference in n. 33).

eye movements—may account for the correlation and, at any rate, need to be eliminated before considering explanations at higher levels.

Binocular combinations in stabilized vision. Fender and Julesz's experiments on the "extension of Panum's fusional area in binocularly stabilized vision" were described above, in the section on examples of the physical model for fusion. In these experiments, relative depth perception persisted in stimuli that had been displaced by 2 deg. on the retina (i.e., the left eye's stimulus moved left by 1 deg. and the right eye's stimulus moved right by 1 deg.). These conditions fall in the middle cell of Table II (secondary NBF, $A \neq B$). The only discrepancy between the prediction of Table II and the observation is that Table II predicts that two images, one of l and one of l', will be seen, whereas visibility of only one was reported. To account for this result, we must carefully examine the procedure.

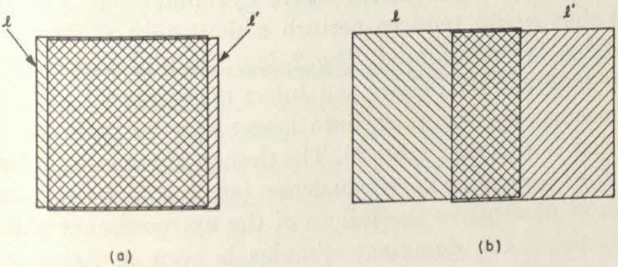


FIG. 18. Cyclopean view of the relative position of the stimulus to the left (l) and right (l') retina in the stabilized-vision experiment of Fender and Julesz (n. 4). The width of the stimuli is 3.4 deg. The orientation (b) illustrates the point at which singleness of vision is lost; (a) illustrates the point at which fusion reoccurs.

Figure 18a illustrates the stimulus conditions at the beginning of a trial, and Figure 18b shows them at the point where singleness of vision was lost. If the stimulus of Figure 18b were presented by itself, both l and l' would be seen, yet when the stimulus of Figure 18a was transformed slowly into that of Figure 18b, subjects saw only one square, an effect which Fender and Julesz explained in terms of an extended fusion area in stabilized vision. Both results can be explained without introducing any new concepts: the crucial detail is Fender and Julesz's inadvertently cunning use of binocular rivalry. For the stimulus of Figure 18a, binocular rivalry determines which of l or l' is seen. Once the lateral disparity between l and l' is about $\frac{1}{4}$ deg. (top-to-bottom depth of the primary NBF), binocular correspondence no longer can be computed within the primary NBF. The active level of the primary NBF is the center level; wherever l and l' conflict on that level, rivalry determines which is suppressed. Because of the dense detail of l and l', it is unlikely that any contour from l can fall against a neutral area in l', or conversely. Contours fall against contours, so that when l' is suppressed, for example, there are no stable islands where contours of l' can break the suppression by l. However, depth relations between portions of l and l' persist because they are computed in the secondary NBF.

As the lateral disparity between l and l' is increased, the suppressed stimulus remains suppressed even after it no longer is competing against contours. This phenomenon can be observed in normal vision, although the model predicts it would be more readily produced in stabilized vision where eye movements and blinks—events that would tend to perturb a delicately achieved suppression state—are eliminated. The 2-deg. limit of single vision (not fusion) observed by Fender and Julesz measures the limit at which the binocular correspondence no longer can be maintained in the secondary NBF (see Figure 5). The change of state in the secondary NBF from binocular correspondence (at $\Delta u = 2$ deg.) to rivalry (at $\Delta u = 0$) triggers the failure of the extraordinary suppression. At this point, the dominant stimulus is seen in its entirety; the other reappears in the regions where it stands alone (Figure 18b). Both stimuli are seen at loci appropriate to their retinal locations.

In brief, the depth that results with binocularly stabilized images is accounted for by depth signals from the secondary NBF, and the singleness of vision (apparent fusion) is accounted for by binocular rivalry. No new principles are involved.[35] In fact, similar suppression phenomena can be observed in normal, unstabilized vision by displacing stimuli laterally beyond the limits at which the eyes diverge to verge on them. In unstabilized vision, binocular rivalry is quite apparent, since only partial suppression is normally achieved.

---

[35] Professor Fender (personal communication) himself has often observed the sudden loss of apparent fusion at the critical separation of binocularly stabilized images. He states that the sensation is one of sudden 'jumping apart' of two images from a single source. The proposed suppression hypothesis, he thinks, would predict that at the instant when suppression fails, the suppressed image merely reappears without any 'jumping apart.' The observation of 'jumping apart' and the assumption that suppression precludes 'jumping apart' causes him to reject the suppression hypothesis. To support the hypothesis it is proposed that the loss of binocular correspondence in the secondary NBF, a loss which is coincident with reemergence of the suppressed stimulus, produces a sensation of 'jumping apart.'

Other functions in the NBF. So far, only two functions of the input have been considered: coarse detail and fine detail. Both of these functions can be accomplished by center-surround types of receptive fields of the appropriate size. Undoubtedly, many other functions of the input are computed. For example, the well-known red-green antagonism in binocular vision suggests the existence of a function pair for green (F) and red ($\sim F$). So far too, it has been assumed that the fine-detail and coarse-detail functions of the primary and secondary NBFs are noninteracting and that the fine-detail system of the primary NBF controls other inputs, such as color and brightness. At present, however, the relations to each other of different function systems—even the specification of such systems—must be regarded as an open question for future theory and experiments.

Two processes that have erroneously been attributed to the NBFs are a change in the mapping of corresponding points with asymmetric vergence (to account for stability of the apparent normal plane) and a change in the receptive field size with convergence (to account for changes in apparent size).[36] In fact, corresponding points, as measured by nonius lines are so remarkably stable[37] that remapping is excluded, and we must look elsewhere for explanations of these perceptual phenomena. The apparent location, orientation, and size of objects depend in a complex way on the interpretation by subsequent analyzers of information available from the NBFs *and other sources* rather than on any restructuring of the NBFs themselves.

Evolution of Binocular Depth Discrimination

Some of the interactions between accommodation, vergence, and fusion are best understood in terms of their evolutionary development. Of the three, according to Walls, accommodation was evolved first and most universally.[38] The most versatile of the accommodation mechanisms is that for comparing blur in different retinal areas. It requires the object whose depth is in question to overlap at least two predetermined retinal areas merely to make a determination of closer or further than the current value of accommodation; even then, the determination of blur is highly dependent on the number of contours within each area. Furthermore, when pupil diameter is reduced in bright daylight, blur becomes an extremely insensitive indicator of depth because of increased *depth of focus*. Therefore, an accurate accommodative depth system would ultimately depend on movements of the lens and of the eye to track depth, and would use the accommodation values as the depth estimate. Unfortunately, systems that require both the eye and the lens to track are too inefficient, too insensitive (relative to the systems described below), and too slow—and therefore seem not to have evolved. Accommodation thus failed to develop into a depth-discrimination mechanism, and depth determination fell to other systems.

---

[36] W. Richards, Spatial remapping in the primate visual system, *Kybernetik*, 4, 1968, 145–156.

[37] Ogle (n. 8).

Head movements provide extremely good cues for the monocular discrimination of depth. The two successive images to the one eye correspond exactly to the two different views of different eyes. The evolution of a depth mechanism based on head movements would suffice for animals that are not endangered by having to move their heads and for animals that capture food with their heads by moving toward it until they encounter it. But animals that use paws, hands, or trunks to capture food need *binocular* depth perception to avoid avoidable trial-and-error reaching. (In fact, such animals have the most binocular overlap in their monocular fields—an indication of dependence on binocular vision.)

The evolution of stereoscopic depth discrimination required, first, overlapping visual fields (which originally evolved to avoid blind spots); second, a binocular registration mechanism (the NBF, which originally evolved to deal with the binocular overlap); and ultimately, vergence movements. To show how these systems evolved together, consider the following example. Suppose a species has evolved an NBF that is ten levels thick, and that an animal can discriminate two depths if their levels in the NBF differ by one level. If the total NBF thickness of ten levels corresponded to 15 deg., the animal could discriminate two depths that differed by 1.5 deg.; if the total depth range of the NBF were 5 deg., the animal could discriminate .5 deg. As the NBF evolved toward the limits of visual acuity and became capable of ever-finer discrimination, it also covered an ever-smaller depth range. The principle is the same as in a microscope: the higher the power, the smaller the range of view. A small NBF depth range implies a critical dependence on accurate centering in the NBF. Therefore, to evolve a sensitive NBF, precise vergence movements were necessary.

---

[38] Walls (n. 13), 247 ff.

A small NBF range also means that vergence is not computed for objects outside this range. At this point, the link of accommodation to vergence becomes an important supplement to the NBF. Thus the development of sensitive binocular depth discrimination required the concomitant development of accurate vergence (to center objects in the NBF) and of linked accommodation (to supplement the range of the NBF).

Conclusion

The neural model represents only a tiny subject of the functions carried out in the visual system and does so in terms of neural components compatible with our present-day knowledge of neural processes. It is not unique: similar functions could be carried out in other ways. Nor does the gometric arrangement of parts of the model—multilayered NBFs—necessarily correspond to neuroanatomy; the model is an hypothesis only about topology (i.e., connectivity).

For all its limitations, the model has one characteristic that is of very general interest. Namely, one function performed by the BCDNs—limiting activity to a single level of the NBF—is representative of a universal kind of function in the nervous system. The BCDNs perform an extreme filtering action that restricts the whole range of potential NBF outputs to output from only one level; it makes only some very selected aspect of the stimulus available for further processing. Since it is inefficient for a system to decide in advance just which selected aspects of every possible stimulus are to be processed, the selection of input for processing must be under dynamic control; the results of previous processing—the previous outputs—must control the future inputs. Therefore, extreme filtering invariably is accompanied by feedback control of the filter; these twin processes produce the phenomena of multiple stable states.

For example, the lens of the eye can be focused exactly on but a single depth plane, an input restriction that places greater weight on the depth plane accommodated and, together with the feedback control of accommodation, leads to the multiple stable states of accommodation. Similarly, the vergence angle between the eyes determines the input from which binocular image disparity is computed, and in turn image disparity controls vergence. Again, filtering and feedback control lead to multiple stable states. Of course, vergence movements are just a minor subclass of eye movements. In general, the direction in which the eyes are pointed determines the input they receive, and this input ultimately determines the direction in which they point. Eye movements represent the ultimate in peripheral filtering under feedback control—a truly multistable system.

The BCDN interactions in the NBF are especially interesting because the filtering is not mechanical but neural. When a binocular correspondence is detected at one level of the NBF, the level becomes active and output from the NBF is restricted to that one level (is filtered). The choice of level depends on responses to previous inputs (is feedback controlled), and the result is multiple stable states of fusion. Other examples of visual processes that exhibit extreme multistability, presumably because of feedback-controlled neural filtering, are binocular rivalry (seeing the stimulus from one eye in a region of the visual field precludes seeing a conflicting stimulus from the other eye in that region) and figure-ground phenomena (seeing an ambiguous figure in one mode precludes simultaneously seeing other modes; seeing a particular figure-ground relation, as on a map, precludes simultaneously seeing other figure-ground relations). Even elementary interactions between successive visual stimuli can be interpreted as exhibiting this multistable 'either-or' characteristic.[39]

In nature, there normally is no reason for organisms simultaneously to execute two antagonistic responses, and organisms cannot do it. Therefore, there is no reason to preserve information that ultimately would lead to antagonistic responses. The neural binocular field is an early locus for the rejection of antagonistic information; many of its paradoxical properties follow therefrom.

---

[39] G. Sperling, Bistable aspects of monocular vision, *J. opt. Soc. Amer.*, 50, 1960, 1140–1141.

APPENDIX A

PROOF THAT BLURRING INCREASES IMAGE-DEFOCUS ENERGY (h_a): THE RELATION OF h_a TO RECEPTIVE-FIELD STRUCTURE

Let $l(x, y)$ be any luminance distribution such that $l(x, y) \geq 0$. Let $b(x, y)$ be any blur function such that $b(x, y) \geq 0$ and $B = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b(x, y)\, dx\, dy = 1$. When $l(x, y)$ is blurred by the function $b(x, y)$ the resulting blurred image $[l_b(x, y)]$ is given by

$$l_b(x, y) = \int_{-\infty}^{\infty} \int_{\infty}^{\infty} l(x', y')b(x - x', y - y')\, dx'\, dy'. \qquad \text{[A-1]}$$

The two-dimensional Fourier transform of $l(x, y)$ is

$$L(\omega, \sigma) = \int_{\infty}^{\infty} \int_{\infty}^{\infty} e^{j\omega x + j\sigma y} l(x, y)\, dx\, dy.$$

The Fourier transform $\mathcal{B}(\omega, \sigma)$ of $b(x, y)$ is defined similarly. Let ∇^i represent the generalized Laplacian $[(\partial^i/\partial^i x) + (\partial^i/\partial y^i)]$. Then Parseval's theorem gives the following equality, which also defines $G(l, i)$:

$$G(l, i) = \iint \left(\nabla^i l(x, y)\right)^2 dx\, dy = \iint (\omega^i + \sigma^i)^2 |\, L(\omega, \sigma)|^2\, d\omega\, d\sigma\,. \qquad \text{[A-2]}$$

$G(l_b, i)$ gives a similar relation in terms of $|L_b(\omega, \sigma)|$, where $L_b(\omega, \sigma) = L(\omega, \sigma)\mathcal{B}(\omega, \sigma)$. Proof of the theorem that blurring increases h_a follows immediately from Equation A-2 if we can show that $|L_b(\omega, \sigma)|^2 \leq |L(\omega, \sigma)|^2$; to show this it is necessary only to show that $|\mathcal{B}(\omega, \sigma)|^2 \leq 1$, since

$$|L_b|^2 = |L|^2\, |\mathcal{B}|^2. \qquad \text{[A-3]}$$

Using E to denote expectation, and recognizing that b is a probability distribution function, gives

$$|\mathcal{B}(\omega, \sigma)| \equiv |E(e^{j\omega x + j\sigma y})| \leq E\, |e^{j\omega x + j\sigma y}| = E\, |1| = 1. \qquad \text{[A-4]}$$

In fact, the middle equality occurs only when $b(x, y) = \delta(x - a_1)\, \delta(y - a_2)$; that is, when there is no blurring, merely exact reproduction or translation.

Equations A-3 and A-4 in conjunction with Equation A-2 give

$$G(l_b, i) \leq G(l, i). \qquad \text{[A-5]}$$

In the text, h_a was defined as $-G(l,\ 2)$. As Equation A-5 unambiguously states, blurring l to produce l_b can only decrease G or, equivalently, increase h_a. (The proof of Equation A-5 is due to M. Sondhi.)

Receptive fields. The definition of h_a in terms of ∇^2 is perfectly logical in terms of receptive fields. The antagonistic center-surround type of receptive field corresponds to a 'blurred' ∇^2.[a] That is, if the center portion of the receptive field were infinitely narrow, and the surround were an infinitely small surrounding annulus which subtracted from the center, the receptive field would correspond exactly to ∇^2. In fact, the center and surround portions of the receptive field have a finite extent; this extent corresponds to a particular blurring (b) of the output of a network of receptive fields. The transformation performed by the receptive field is thus approximated by $b\nabla^2$. Because differentiation (∇^i) commutes with convolution (blurring), the finite center-surround dimensions of the receptive fields can be treated as blur in the stimulus, and the fields themselves considered to be perfect differentiators ($\nabla^2 b$). The defocus of the lens adds additional blur to the image; it is this variable source of blur, h_a, which is the stimulus for accommodation.

The use of the square $\nabla^i l(x,\ y)^2$ in Equation 8 of the text and Equation A-2 is for mathematical elegance. The replacement of $(\nabla l)^2$ by $|\nabla l|$ might perhaps bring the expression closer to physiological plausibility (separate 'on' and 'off' channels convey the positive and negative values of ∇l), but it would not materially alter the main results (Equation A-5). That is, when the absolute value replaces the square, Equation A-5 becomes

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\nabla^i l_b(x,\ y)|\, dx\, dy \le \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\nabla^i l(x,\ y)|\, dx\, dy \qquad \text{[A-6]}$$

where l_b is a blurred image of l.

In Equation A-6, strict equality holds for $i = 0$; equality may still hold even when $i \geq 1$ for certain combinations of stimuli (l) and blur functions (b). The condition for inequality is that blurring cause some light from an area where $\nabla^i l > 0$ to spill into an area where $\nabla^i l < 0$ (or vice versa). For $i \geq 2$, this spillage always occurs in physically realizable images with nontrivial blur functions, and so the absolue value, as well as the square, could serve in the evaluation of blur. Finally, as a practical matter, for special $l(x, y)$ which are themselves very 'blurred,' the additional effect of defocus may be negligible; in these cases, the theory predicts, it will be impossible to accommodate the images.

---

[a] For an early insight into this problem, see E. Mach, Ueber die physiologische Wirkung räumlich vertheilter Lichtreize, IV, *Sitzungsber. math.-naturwis. Classe Kaiserl. Akad. Wissensch.*, 57 (1868), No. 2, 11–19.

APPENDIX B

ANALYSIS OF NEURON-PAIR INTERACTIONS

In terms of its function, a neuron is defined by its input-output relations; these are approximated by a nonlinear differential equation of the form

$$dy(t)/dt + (A + B(x) + C(y))y(t) = x(t). \qquad \text{[B-1]}$$

Here, $x(t)$ is the input, $y(t)$ is the output; A is a constant, B is a function of the entire input (feedforward inhibition), and C is a function of the entire output of all the neurons at the same level (feedback inhibition).[b]

Elementary neuron pair. Consider the pair of neurons illustrated in Figure B-1a. These neurons interact only by mutually inhibiting each other; the constants (k) of mutual inhibition are assumed to

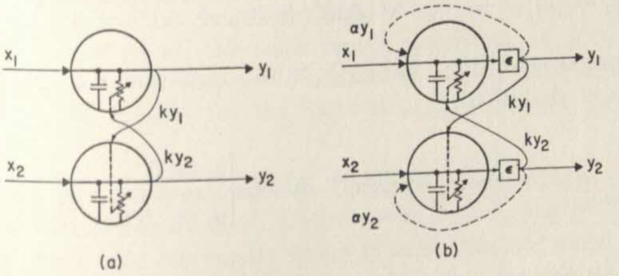


FIG. B-1. Reciprocally interconnected neurons. (a) Elementary neuron pair: each neuron consists of an RC stage; the value of R is controlled by the output of the neighboring neuron. (b) Neurons with positive feedback: threshold, ϵ, for output has been added, and there is an excitatory feedback loop, αy_i, from the output back into the input.

---

[b] To relate Equation B-1 to cellular parameters, consult W. Rall, Theoretical significance of dendritic trees for neuronal input-output relations, in R. F. Reiss (ed.), *Neural Theory and Modeling*, 1964; G. Sperling and M. M. Sondhi, Model for visual luminance discrimination and flicker detection, *J. opt. Soc. Amer.*, 58, 1968, 1133–1145.

be equal; and there is assumed to be no threshold and no self-inhibition. Thus the term $B(x) = 0$; $C(y) = ky$; the dimensional constant (A) is taken to be 1.0. The analysis becomes quite simple if it is restricted to steady-state outputs ($dy/dt = 0$) in response to step inputs [$x(t) = \text{constant}, t > 0$]. In fact, very little generality is lost thereby. For such a neuron pair, Equation B-1 reduces to

$$y_1 = \frac{x_1}{1 + ky_2} \qquad \text{[B-2]}$$

and

$$y_2 = \frac{x_2}{1 + ky_1}. \qquad \text{[B-3]}$$

The inputs (x_i) are restricted to positive values and labeled so that $x_2 \geq x_1 \geq 0$. Define Δy as $y_2 - y_1$ and Δx as $x_2 - x_1$. Multiplying out Equations B-2 and B-3, and subtracting, gives

$$\Delta y = \Delta x. \qquad \text{[B-4]}$$

This result holds for all values of k.

Adding Equations B-2 and B-3, and reducing the result algebraically, gives

$$(y_1 + y_2) = -\frac{1}{k} + \sqrt{(\Delta x)^2 + 2(x_1 + x_2)/k + 1/k^2}. \qquad \text{[B-5]}$$

For special conditions (when $k \gg 0$), Equation B-5 can be substantially simplified:

$$k\Delta x \gg 2(x_1 + x_2) \Rightarrow \begin{cases} y_1 \approx 0 \\ y_2 \approx \Delta x \end{cases} \qquad \text{[B-6]}$$

and

$$k\Delta x \ll 2(x_1 + x_2) \Rightarrow \{y_1 \approx y_2 \approx \sqrt{x/k}. \qquad \text{[B-7]}$$

Equation B-6 holds for $x_1 \neq x_2$ and Equation B-7 holds for $x_1 \approx x_2$. These results show that the elementary neuron pair is a system for performing subtraction. The system is all the more remarkable because the interaction is not subtractive but divisive (the inhibition terms appear in the denominators of Equations B-2 and B-3). The output Δy always equals the input Δx, and when k is large, the smaller output is approximately zero. For large k, at most one of two neurons is active, the one with the larger input, and this neuron's output is approximately Δx.

Elementary neuron N-tuple. When more than two neurons interact mutually in the same way as the neuron pair (i.e., as a neuron N-tuple), Equations B-2 and B-3 generalize to

$$y_i = \frac{x_i}{1 + k \sum_{j \neq i} y_j}. \qquad \text{[B-8]}$$

For large k and N the following properties are derived from Equation B-8. Let x_1 be the largest input and let all other inputs (x_i) be equal. Two cases are considered: first, when $x_1 > \Sigma_2{}^n\, x_i$ (see Equation B-9) and second, when $x_1 \leq \Sigma_2{}^n\, x_i$ (see Equation B-10).

$$x_1 = Nx_2 + \gamma \text{ and } \gamma > 0 \Rightarrow \begin{cases} y_1 \approx \gamma \\ y_j \approx \dfrac{x_2}{k\gamma} \approx 0 \end{cases} \qquad \text{[B-9]}$$

$$x_1 = Nx_2 \Rightarrow \begin{cases} y_1 = x_1{}^{2/3} k^{-1/3} \rightarrow 0 \\ y_j = x^{1/3} k^{-2/3} N^{-1} \rightarrow 0 \end{cases} \qquad \text{[B-10]}$$

When the largest input (x_1) is less than the sum of other inputs, the output also yields a power of k ($k^{1/2}$) in the denominator, much as in Equation B-10. Thus, the neuron N-tuple retains the desirable features of the elementary pair only when one neuron has a larger input than all the other inputs combined. Otherwise, all outputs depend on k and go to zero as $k \rightarrow \infty$. (The author is indebted to M. Sondhi for the derivation of Equations B-9 and B-10.)

Neurons with positive feedback. Consider an elementary neuron pair, like that of Figure B-1b. Each neuron feeds a part of its output, $\alpha\, y$, back into its input, where it acts like an added input signal. (The threshold is ignored, temporarily.) The equation describing this modified pair is

$$y_i = \frac{x_i + \alpha y_i}{1 + k y_j}. \qquad \text{[B-11]}$$

By defining β as $1/(1 - \alpha)$ for $\alpha < 1$, it is a simple matter of algebraic manipulation to reduce Equation B-11 to

$$y_i = \frac{(\beta x_i)}{1 + (\beta k) y_j} = \frac{x'}{1 + k' y_j}. \qquad \text{[B-12]}$$

The right-hand term of Equation B-12 shows that by redefining input as $x' = \beta x$ and the inhibitory coefficient as $k' = \beta k$, the basic form of Equation B-2 remains unchanged by positive feedback. The net effect of positive feedback is to multiply k and Δx by β. This is no small effect: as α approaches 1, β approaches ∞; for $\alpha \geq 1$, β is infinite. Thus, a relatively small amount of positive feedback can enormously increase the strength of interaction (k) between neurons, and their gain (β).

Effect of threshold (ϵ). Neuronal thresholds arise naturally from the two-stage conversion of input to output: a neuron's input is converted, first, into an internal voltage and, second, into an output spike rate. Spike rates typically are proportional to internal voltage, but as rates cannot fall below zero, there is a threshold voltage. If the internal resting voltage of a neuron (zero input) is ϵ below its threshold-for-firing voltage, then inputs will not be reflected in the output until they perturb the resting voltage by ϵ (Figure B-1b). The output of a neuron with a threshold and positive feedback is given by

$$y_i = \max\left[\frac{x_i + \alpha y_i}{1 + k \sum_{j \neq i} y_j} - \epsilon, 0\right]. \qquad \text{[B-13]}$$

Rather than develop the full complexities of Equation B-13, only one simple case is here considered, the case where $\alpha \geq 1$. Consider first a neuron acting in isolation: $x_i = 0$, $i \neq j$. (Note that $x_i = 0$ implies $y_i = 0$.) Because $\alpha \geq 1$, any input $x_i > \epsilon$ causes $y_i \to \infty$, or more realistically, $y_i \to Y$, the maximum possible output. A neuron (i) acting in isolation is either fully turned on ($y_i = Y$) or silent ($y_i = 0$) depending on whether $x > \epsilon$ or $x \leq \epsilon$.

Monoactivity in an interacting system of neurons with thresholds and positive feedback. Let X be the maximum possible input; let Y be the maximum possible output; let $\alpha \geq 1$. Consider first what happens when all inputs except x_i are zero; that is, $x_j = 0$, $j \neq i$. This is the same case as neuron i acting in isolation: $y_j = 0$ (for $j \neq i$) and either $y_i = Y$ (for $x_i > \epsilon$) or $y_i = 0$ (for $x_i \leq \epsilon$).

When more than one neuron receives nonzero inputs, the first neuron (i) whose input (x_i) exceeds ϵ turns fully on ($y_i = Y$). The condition for monoactivity of a network of N-neurons is quite simple; it is that any active neuron has sufficient output to prevent all other neurons from turning on. For example, if the first neuron (i) whose input (x_i) exceeds ϵ were to produce an output (Y) of sufficient magnitude to keep all the others silent, this would be a highly stable state. It would persist until the input changed. When x_i finally subsided to ϵ, another neuron could take over; it would silence the others, and so on. Sufficient conditions for this kind of monostable interaction are developed below.

Neuron j is maintained in perpetual silence by neuron i when

$$\frac{x_j}{1 + kY} \leq \epsilon. \qquad \text{[B-14]}$$

(Note that the denominator of Equation B-14 contains kY, and not βkY as suggested by Equation B-12, because it is assumed that Y already is the largest value of y_i. The numerator of Equation B-14 does not contain β because it is assumed that neuron j's threshold is not exceeded; therefore there is no positive feedback.) Rearranging Equation B-15 gives

$$(1 + kY) \geq x_j/\epsilon = m. \qquad \text{[B-15]}$$

The left side of Equation B-15 is the total inhibition of neuron j (by neuron i); physiologically, it represents the factor by which neuron i must change the effective conductance of neuron j to silence it. The factor $m = x_i/\epsilon$ is the number of times by which x_j exceeds the threshold (ϵ). Let $X/\epsilon = M$. To guarantee monostability, the effective neuronal conductance change (left side of Equation B-15) must exceed the largest m, giving $(1 + kY) > M$. When this condition is satisfied, neuron x_j is silenced by x_i even when x_j is receiving its maximum possible input.

The value of M depends largely on ϵ; when ϵ is large, M is small and easily achieved. To avoid the possibility of having no active neurons when ϵ is large, it can be assumed that some neurons (i.e., neurons of the middle level of the NBF) receive spontaneously active inputs; these effectively reduce the ϵ requirement without changing M.

To obtain bounds on M, consider the two factors which determine the effective change in neuronal shunting conductance (the left-hand term of Equation B-15). Let the factor $f_1 = (1 + kY)$ be the direct increase in conductance wrought by inhibition. When the site of excitation is separated from the site of inhibition, the effectiveness of the inhibition can be increased; the increased effectiveness is represented by the factor f_2. Assuming the attainable values of f_1 and of f_2 are about 3 gives a net $f_1 f_2$ of about 9ϵ for the range of opposing inputs that can be totally silenced by an active neuron. For inputs ranging up to 9ϵ, this system is monostable and unbiased. If one or more inputs exceed ϵ, one (and only one) neuron is active; otherwise, all are silent. This system has the properties assumed for it in the neural model.

THE EFFECT OF THE SHAPE OF DISTRIBUTION OF RESPONSE STRENGTH ON GUESSING ERRORS IN PROBLEM SOLVING

Carl P. Duncan, *Northwestern University*

Abstract. Subjects were given problems each consisting of a stimulus word with its five, seven, or nine strongest word associates. They attempted to guess the middle-strength (target) associate. Since the hierarchy of strength between a stimulus and its associate responses is such that differences in strength among associates become progressively smaller as strength between stimulus and response decreases, associates weaker than the target associate are more similar in strength to the target than are associates stronger than the target. It was therefore expected that in trying to guess the target associate, more errors would be made to weaker than to stronger associates. This was found in all conditions. More generally, it appeared that the farther removed in strength—whether stronger or weaker—an associate was from the target associate, the less it was interfering. Giving the subjects a cue to associates stronger or weaker than the target word had some effect on, but in no way eliminated, the strong tendency to make more errors on weaker associates.

In three previous studies it has been found that when subjects attempted to solve problems that required guessing the *n*th item in a verbal-response hierarchy of five members, items of intermediate habit strength were more difficult to guess than either stronger or weaker members of the hierarchy. The problems used have been (*a*) stimulus words each with its five strongest word associates,[1] (*b*) sets of five bigrams where the bigrams within a set were of different frequencies of occurrence in words,[2] and (*c*) sets of five words, the words in a set having different frequencies of occurrence in the printed language.[3] In all of these studies, interest was cen-

---

Received for publication June 11, 1970. The study was supported in part by Grant HD-00901 from the Institute of Child Health and Human Development, National Institutes of Health, United States Public Health Service. Thanks are due Donald Lehr and Robert Koch.

[1] C. P. Duncan, Response hierarchies in problem solving, in C. P. Duncan (ed.), *Thinking: Current Experimental Studies*, 1967, 18–41.

[2] R. L. Dominowski, Problem difficulty as a function of relative frequency of correct responses, *Psychon. Sci.*, 3, 1965, 417–418.

[3] N. H. Jurca and C. P. Duncan, Problem solving within a word-frequency hierarchy, *J. verb. Learn. verb. Behav.*, 8, 1969, 229–233.

tered primarily on the difficulty of guessing the correct response as a function of its strength (i.e., its position in the hierarchy). No systematic examination was made of guessing errors. One study did present graphs of error responses, but without analysis (n. 1).

If the members of a five-item hierarchy of strength are numbered from 1 (strongest) to 5 (weakest), then, as noted, Item 3 has usually been the most difficult to guess. This, plus the fact that the hierarchy contains both stronger and weaker responses on which errors can be (and are) made, makes the middle position particularly useful for studying guessing errors. The present experiment was an examination of guessing errors around the middle-strength item in hierarchies of varying numbers of items.

The hierarchies that were used in the present experiment were taken from the overall hierarchy that is shown in Figure 1. The figure was derived from Duncan and Wood's study in which 500 subjects each gave five word associates to each of 20 stimulus words taken from Russell and Jenkins.[4] For each stimulus word, the frequency of occurrence of each of the 10 associates most often

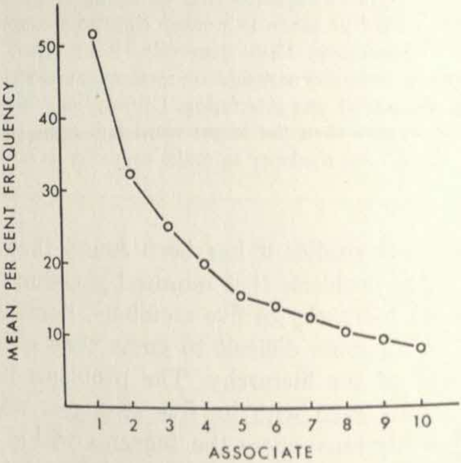


FIG. 1. Mean percent strength (frequency of response) of each of the first 10 associates to 20 stimulus words (data from Duncan and Wood [n. 4.])

---

[4] C. P. Duncan and G. Wood, Norms for successive word associations, *Psychon. Monogr. Suppl.*, 1, 1966, No. 7, 203–206; W. A. Russell and J. J. Jenkins, The complete Minnesota norms for responses to 100 words from the Kent-Rosanoff word association test, Office of Naval Research, 1954.

given to that stimulus word was recorded. For the 20 responses, one per stimulus, identified as Associate 1 (highest strength, most frequently given), the mean frequency over the 20 stimuli was computed, then converted to a percentage based on 500 subjects. The same was done for each of the other associates (Associates 2–10) for each stimulus. The mean percent frequencies are plotted for each of the 10 associates in Figure 1, where the x-axis indicates, from left to right, increasing weakness of association.

Suppose subjects are given problems, each consisting of a stimulus word with its five strongest associates, taken from the study on which Figure 1 is based, and are asked to guess which is Associate 3 in all problems. The plotted points for Associates 1–5 in Figure 1 are assumed to represent the relative strengths of the five associates of the 'mean problem' (of all problems) on which the subjects have to work. Response-hierarchy theory as applied to problem solving seems to imply that in trying to discover the solution response to a problem, responses higher in the hierarchy (greater in strength) tend to be elicited by the problem stimuli and thus delay or block solution.[5] However, another possibility is that responses more similar in strength to the solution response are more interfering. In a negatively accelerated hierarchy such as the one shown in Figure 1, this would mean that weaker associates might be more interfering than stronger associates. Thus, in trying to guess Associate 3, Associates 4 and 5 might be more interfering than Associates 1 and 2, since the weaker two are closer in strength to Associate 3 than are the stronger two. The same considerations should apply anywhere in the hierarchy, or at least down to asymptotically low levels of strength.

The purpose of the present study was to compare errors made to responses both stronger and weaker than a target response in the hierarchy of Figure 1, and to do so at different absolute levels in the hierarchy. A second purpose was to determine whether patterns of guessing errors would be differentially influenced if cues indicating stronger or weaker responses were given to the subjects.

METHOD

Design. The design was a 3 × 3 factorial. The number variable was the number of associate words presented with a stimulus word. For each of the 20 stimulus words, from the study of which Figure 1 is based, the first five, first seven, or first nine most-frequently-given associates were used for different groups of *S*s. The mean strength of each of these nine associates, over 20 problems, is shown in Figure 1. In each number condition, *S*s attempted to guess, on each of the 20 problems, the middle-strength associate: Associate 3 in Group 5, Associate 4 in Group 7, and Associate 5 in Group 9; thus, these are the target associates.

---

[5] Duncan (n. 1); I. Maltzman, Thinking: From a behavioristic point of view, *Psychol. Rev.*, 52, 1955, 275–286.

The other variable involved cuing. In the no-cue condition, *S*s were given only the general instructions (see below). In the higher-cue group, *E* told *S*, as each problem was presented, which associate was next higher in strength than the target associate. Thus, Groups 5H, 7H, or 9H were told which word was Associate 2, 3, or 4 respectively. For the lower-cue *S*s, in Groups 5L, 7L, or 9L, *E* told *S* the word with the next lower strength than the target, namely, Associate 4, 5, or 6 respectively.

Each of the 20 problems was typed on a separate 5 × 8 in. card. The stimulus word appeared near the left margin of the card and was followed by a colon. The five, seven, or nine associate words were in a horizontal row to the right of the stimulus word. The associates were arranged in three different left-to-right orders, so there were three sets of cards for each number of associates, each set used for one-third of the *S*s in a group.

Procedure. All *S*s were first given a brief description of the obtaining of college students' word associations to stimulus words. Strength of association was defined as frequency, that is, as the number of students all giving the same associate. As an example, *S*s were told that to the stimulus word *forest* (not one of the 20 used here), the largest number of students responded *green,* so it was the first or strongest associate; that a smaller number gave *dark,* so it was the second strongest associate; that an even smaller number gave *cool,* so it was the third strongest associate; and so forth. After *E* felt sure *S* understood, *E* said, "OK, now I am going to show you a card with a stimulus word on the left. The five [seven, nine] strongest associates that students gave to that word are printed in a random order on the right. I want you to guess what you think was the third [fourth, fifth] strongest associate." Then to the cued groups, *E* said, "To help you I will tell you the associate next stronger [weaker] than the one you are to guess, that is, I will tell you the second [third, fourth, or fourth, fifth, sixth] and you are to guess the third [fourth, fifth]." For *S*s in cued groups, *E* gave the cue word just after *S* was handed each card.

On each problem *S* had to keep guessing until he identified the target associate. The *E* recorded which associate *S* chose on every guess.

Subjects. The *S*s were 270 students from introductory psychology courses, 30 in each of the nine groups.

RESULTS

For each subject, guesses (errors) to all associates stronger than the target associate were summed over the 20 problems. Since the number of stronger associates varied from one in Group 5S (given Associate 2, leaving only Associate 1 as a stronger associate on which to make errors) to four in Group 9N (no cue), the error sums were divided by the appropriate number of associates to give a mean error score to stronger associates. The same was done on guessing errors to associates weaker than the target associate. Thus, there were two error scores for each subject. Table I shows the mean of these error scores over the 30 subjects in each of the nine groups. In the table, the columns headed by S contain the means for errors to stronger associates, W indicates weaker error means. It can be seen that the W mean was numerically larger than the S mean in each of the nine groups.

TABLE I

MEAN ERRORS TO STRONGER (S) AND TO WEAKER (W) ASSOCIATES IN ALL NUMBER AND CUE CONDITIONS

	Number of Associates					
	5		7		9	
Cue	S	W	S	W	S	W
None	9.37	10.75	9.31	10.32	8.59	9.41
Higher	8.47	8.92	8.63	9.28	9.08	10.33
Lower	8.90	11.33	9.00	10.50	8.64	9.84

Analysis of variance of the means showed that the comparison of major interest (S versus W means in Table I) was highly significant [$F\ (1, 261) = 101.80, p < .001$]. Associates weaker than the target associate clearly attracted more errors. Also, the S-W variable showed a significant interaction with cue [$F\ (2, 261) = 4.98$, $p < .01$]. Other interactions with the S-W variable were not significant.

The effect of the cue variable was significant [$F\ (2, 261) = 3.25$, $p < .05$]; the higher-cue subjects made slightly fewer total errors (S and W combined) than either no-cue or lower-cue groups. The number variable did not significantly influence total errors. However, cue and number yielded a significant interaction [$F\ (4, 261) = 21.24$, $p < .01$]. In part this is because the higher-cue groups made fewer total errors than the no-cue or lower-cue groups with five or seven, but not with nine, associates.

It is clear that associates weaker than the target associate attracted more errors. If, as suggested earlier, this is because weaker associates are more similar in strength to the target word than are stronger associates, then *any* associate, stronger or weaker, that is closer in strength to the target should produce more errors than an associate farther removed in strength. This possibility was examined by plotting Figure 2. The figure shows mean responses to each associate separately (including the three target words; Associates 3, 4, and 5) for the three groups given no cues. (Thus all associates were available to be guessed.) First, it may be noted that, as reported, associates weaker than the target word yielded more errors. But the figure also suggests that errors exhibited declining gradients on both sides of a target word: the farther away in strength from a target, the fewer the errors.

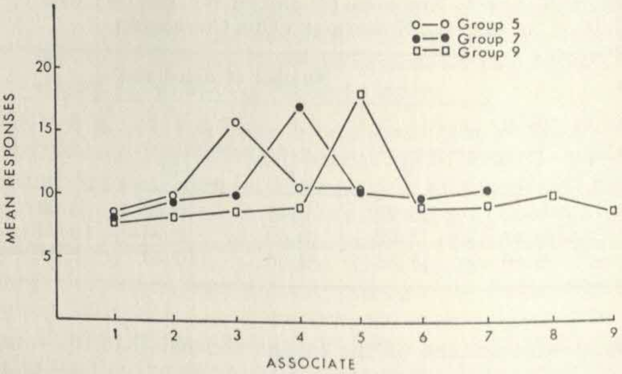


FIG. 2. Mean responses to each associate in groups not given cues. The highest point on each curve represents correct guesses of the target associate; all other points indicate errors.

A three-variable analysis of variance was done on the data in Figure 2; one variable was the three groups, the second variable was the distance (one step or two steps) separating an error associate from the target associate, and the third variable was strength (stronger or weaker) of an associate compared to the target associate. The significant effect [$F\ (2, 87) = 131.38, p < .001$] for groups was probably the result (as were the differences among the groups in performance on the target word) of differing probabilities of guessing a word by chance across five, seven, or nine words. Of greater interest are the terms for distance and for strength. For distance [$F\ (1, 87) = 3.96, p < .05$], errors were fewer on words two steps versus one step removed from the target. For strength [$F\ (1, 87) = 14.18, p < .001$], as previously reported, more errors were made to weaker associates. The interaction of groups and strength was also significant [$F\ (2, 87) = 6.61$, $p < .005$]; the difference between weaker and stronger associates was less in Group 9 than Group 5. The interaction between distance and strength fell just short of significance [$F\ (1, 87) = 3.61$]; there is a suggestion that the gradient effect was sharper among stronger than among weaker associates, especially where this would be expected, namely, higher in the hierarchy (Group 5).

DISCUSSION

Response-hierarchy theory, mainly derived from Hull, has been applied to thinking—particularly problem solving—by Maltzman, Osgood, Cofer, and others, who assumed that a problem tends to elicit a number of responses and that these responses differ in probability of occurrence.[6] In this respect the potential responses form a hierarchy of varying strengths. The response that is the solution to the problem is at some medium- or low-strength position in the hierarchy and is therefore less likely to occur, at least initially, than stronger responses. It seems to have been implied by the theorists that a problem is difficult of solution to the extent that stronger, high-hierarchy responses interfere with discovery of the weaker solution response.

Interference from responses stronger than the solution certainly occurs. In fact, it is possible that in response-production or response-retrieval problems, where the pool of responses to be searched is not clearly specified for the subject, stronger incorrect responses may provide the major source of interference. Even in response-selection or multiple-choice problems, where the pool of potential responses is specified (and the number of responses usually quite limited), stronger responses do attract errors, as found in this and previous studies (n. 1). But the present experiment shows that in a response-selection problem, it can occur that responses weaker than the solution give the subject more trouble than stronger responses.

This finding probably results from what is assumed to be the general principle: that in a response hierarchy, responses similar in strength tend to interfere with each other more than responses less similar in strength. If such a principle holds, errors should show gradients, declining in numbers as responses are farther removed from a target response. Such gradients appeared to some degree in the present study and in previous studies (nn. 1, 2, 3).

---

[6] C. L. Hull, The concept of the habit-family hierarchy and maze learning, *Psychol. Rev.*, 41, 1934, 33–54, 134–152; Maltzman (n. 5); C. E. Osgood, *Method and Theory in Experimental Psychology*, 1953; C. N. Cofer, Reasoning as an associative process: III, The role of verbal responses in problem solving, *J. gen. Psychol.*, 57, 1957, 55–68.

One of the independent variables in the present study was the number of associates (five, seven, or nine) of the stimulus word. Since no previous study had used more than five responses among which the subject had to guess, the number variable was used partly to achieve greater generality. Use of more than five associates, in problems where the subject always had to guess the middle associate, also made available potential error responses farther removed in strength from the target associate. On the comparison of major interest, namely, errors to stronger versus errors to weaker associates, the number variable, as expected, had no effect; the F value was not significant. The tendency to make more errors to weaker wrong associates did not vary whether there were five, seven, or nine associates to be searched. The number variable also had no differential effect on total errors.

The other independent variable, cue, involved telling the subject the associate just higher (in strength) or just lower than the target word, on each problem. Cue did interact ($p < .01$) with the tendency to make more errors to weaker associates. To examine this, the number variable in Table I was collapsed and the means for the three cue groups, for stronger and weaker associates separately, were computed. These means, with the stronger-associates mean first, were: no-cue group, 9.09 and 10.16; higher-cue group, 8.73 and 9.51; and lower-cue group, 8.85 and 10.56. The difference between stronger and weaker means was 1.07 in the no-cue group. In the higher-cue group the difference dropped to .78, whereas in the lower-cue group the difference increased to 1.71. It might seem that where weaker associates are more interfering, giving the subject a cue by telling him one of the weaker associates would be helpful in reducing the error-attractiveness of weaker associates. However, the subject also knew that the cued associate was one close in strength to the target associate. If he attempted to use this information by looking for associates that seemed to him similar in strength to the cue associate, then for the subjects given a weaker associate as a cue (lower-cue group), the associates more similar in strength to that cue would be still weaker associates

(Figure 1). For the subjects given a stronger associate as a cue (higher-cue group), the associates more similar in strength to the cue were also associates that were weaker than the cue, but one of those was the target associate. (It may also be noted that total errors, stronger and weaker combined, were fewest in the higher-cue group.) Thus, the lower cue was the least rather than the most helpful, perhaps because it attracted the subject's attention to the difficult-to-distinguish weaker associates.

RELEVANT AND IRRELEVANT SALIENCY IN CONCEPT LEARNING

HAROLD D. FISHBEIN, *University of Cincinnati*,
ROBERT C. HAYGOOD, *California State College, Long Beach*,
and DIXIE FRIESON, *Kansas State University*

Abstract. Two concept-learning experiments are reported. The first compared performance under three stimulus configurations: with the two relevant dimensions highly salient; with two irrelevant dimensions highly salient; and with no highly salient dimensions. Performance was best when the relevant dimensions were highly salient, next best when two irrelevant dimensions were highly salient, and poorest when there were no highly salient dimensions. The second experiment explored several possible explanations of these results, which are contrary to existing notions of the effects of irrelevant saliency. Its results suggest that the presence of highly salient irrelevant dimensions improved performance because the subjects examined them in the early trials of the experiment and then could permanently eliminate them from further consideration.

It is now generally accepted that variations in the saliency, vividness, or obviousness, of the stimulus materials have an important influence on human learning.[1] Several different methods have been used in the attempt to manipulate saliency. One common method, called perceptual *isolation*, is to make some aspect of the stimulus material stand out by making it perceptually or conceptually unique. The effects of isolation are commonly identified with the name of von Restorff, who introduced isolation in a serial-learning task by embedding a single number in a list of nine syllables.[2] The number was learned faster and retained better than the same number embedded in a more heterogeneous list containing other numbers, syllables, and nonsense forms.

Received for publication March 9, 1970. This research was supported in part by Research Grants MH-11283 and MH-18249 from the National Institute of Mental Health, United States Public Health Service.

[1] T. Trabasso and G. H. Bower, *Attention in Learning*, 1968; W. P. Wallace, Review of the historical, empirical, and theoretical status of the von Restorff phenomenon, *Psychol. Bull.*, 63, 1965, 410–424.

[2] H. von Restorff, Ueber die Virkung von Bereichsbildung im Spurenfeld, *Psychol. Forsch.*, 18, 1933, 299–342.

In concept learning, the beneficial effect of isolation was demonstrated in the earliest objective studies, those of Hull.[3] Hull produced isolation by printing the relevant elements of the stimulus patterns (Chinese characters) in red, leaving the remainder of the pattern black. Relative to the condition in which all elements were black, performance in the isolation condition was superior.

A second method of manipulating the saliency of a stimulus dimension is to change the *difference* between values or levels of the dimension. For example, in Archer's experiment with geometric designs, the size dimension was made more salient by increasing the difference between the two sizes used, even though neither of the sizes taken individually would be considered more salient.[4] Archer's results showed that increasing the saliency of a relevant dimension led to superior performance, while increasing the saliency of an irrelevant dimension generated poorer performance.

These effects of relevant and irrelevant saliency have been confirmed by other workers. For example, both the Hull and Archer techniques were included in a study by Trabasso.[5] Using flower designs in which the angle of the leaves with the stem was the relevant dimension, Trabasso found that coloring the angle red (as in Hull; see n. 3) and increasing the difference between angles (as in Archer; see n. 4) both improved performance.

The finding that relevant saliency helps performance and irrelevant saliency hurts performance has generally been explained in terms of *attention*. If the relevant dimension is highly salient, the subject's attention is drawn to it, and waste-motion studying and trying to use irrelevant information is reduced. Conversely, if an irrelevant dimension is highly salient, then the subject's attention is drawn away from the relevant information, and learning is delayed. This view has been formalized in a mathematical model for concept learning presented by Trabasso and Bower.[6] The model assumes that subjects solve a concept-learning problem by selecting a subset of one or more of the available stimulus dimensions (cues) and testing that subset until either the correct solution has been identified or an error has been made. Following an error, the subject samples (with replacement, and with local consistency) a new subset of cues to test. The model assigns to each cue a weight representing its saliency, or noticeability. The attention value of a cue is its saliency relative to that of the other cues available, and the relative cue weight represents the probability of sampling that cue. Since the subject must sample the relevant cue to solve the problem, the model makes a straightforward prediction that an increase in the saliency of the relevant cue increases the probability of its being sampled and hence leads to more rapid solution. Conversely, an increase in the saliency of irrelevant cues reduces the relative cue weight of the relevant dimension and thus leads to slower solution.

---

[3] C. L. Hull, Quantitative aspects of the evolution of concepts, *Psychol. Monogr.*, 28, 1920, No. 1 (Whole No. 123).

[4] E. J. Archer, Concept identification as a function of obviousness of relevant and irrelevant information, *J. exp. Psychol.*, 63, 1963, 616–620.

[5] T. Trabasso, Stimulus emphasis and all-or-none learning in concept identification, *J. exp. Psychol.*, 65, 1963, 398–406.

[6] See n. 1 above; see also G. H. Bower and T. Trabasso, Concept identification, in R. C. Atkinson (ed.), *Studies in Mathematical Psychology*, 1964.

It is clear that the performance decrement for increased irrelevant saliency depends in part on the typical inability of subjects to remember past stimulus information and the fate of past hypotheses tested.[7] The reason is that if the subject can remember having tried an irrelevant dimension, the negative effect of having been temporarily misled is countered by the positive effect of having the remainder of the problem simplified by the permanent elimination of that dimension. Thus if the stimulus materials provide subjects some systematic way to remember having tried salient irrelevant dimensions, the performance decrement might be eliminated or even reversed. The purpose of the first experiment reported below was to explore this possibility.

EXPERIMENT I

One obvious method of facilitating permanent recall of irrelevant dimensions already tested is that of perceptual isolation. In a distributed simulus pattern such as an array of typewritten letters, one could substitute pairs of numbers for some of the letter dimensions, much as von Restorff embedded a number in a list of syllables. This form of isolation allows the subject to partition the stimulus set into two subsets, letters and numbers. The studies cited earlier indicate that the subject's attention should be directed to the isolated number dimensions, leading to faster learning—if the number dimensions are relevant—than when all dimensions are letter dimensions. If the number dimensions are irrelevant, however, the subject's ability to remember that they do not work should act to reduce or eliminate the negative effect of irrelevant saliency.

---

[7] T. Trabasso and G. H. Bower, Memory in concept identification, *Psychon. Sci.*, 1, 1964, 133–134; R. C. Calfee, Recall and recognition memory in concept identification, *J. exp. Psychol.*, 81, 1969, 436–440.

Method

Subjects and design. The *S*s were 72 summer-session students at Kansas State University, who were paid for participation. Twelve Ss were assigned to each of six treatment combinations in a stratified random order which ensured that all groups were filled at the same rate. The experimental design was a 3 × 2 factorial with three levels of solution type and two levels of solution spacing. The three solution types were (*a*) two letters, contained in an eight-letter array, (*b*) two letters, contained in a six-letter, two-number array, and (*c*) two numbers, contained in a six-letter, two-number array. The two solution spacings were (*a*) adjacent relevant dimensions and (*b*) relevant dimensions separated by one irrelevant position. Examples of positive instances of the six solutions can be seen in Table I, where the solution in each case is italicized.

TABLE I

EXAMPLES OF POSITIVE INSTANCES FOR EACH OF THE SIX EXPERIMENTAL CONDITIONS: EXPERIMENT I

E	f	g	H	*i*	*J*	k	L
e	F	G	H	*3*	*5*	K	1
E	F	2	5	*i*	*J*	K	L
e	f	G	*h*	I	*J*	k	1
E	f	g	*3*	i	*5*	k	1
e	f	3	*h*	4	*J*	k	L

The spacing variable was included partly for control purposes and partly because an earlier study had demonstrated an effect of grouping by proximity when the stimulus array was physically arranged into pairs of dimensions.[8] The intent was to see if the negative effect of separation of relevant dimensions would also occur when the dimensions were evenly spaced. The result here, like that of Trabasso and Bower (see n. 1), was essentially negative, and little more need be said about the spacing variable.

Task and materials. The *S*s performed two tasks. The first was a simple pretraining task that utilized geometric figures with three binary dimensions and required *S*s to learn a two-dimensional conjunctive solution such as 'red and square.' All *S*s received the same pretraining. For the main experimental task, the stimuli consisted of eight letters, or letters and numbers, typewritten on 3 × 5 in. cards (specific examples in Table I). Each position of the row represented a stimulus dimension, which could take on two values. For positions occupied by letters, the values for each dimension were 'capital' and 'lower case' (e.g., D *or* d). For the two positions occupied by numbers, the two values were 2 or 3, and 4 or 5. In each case there were two relevant and six irrelevant dimensions, with the concept involving either the fourth and sixth or fifth and sixth positions (from the left). Patterns failing to contain both relevant attributes were negative instances, and the proportions of positive and negative instances were held constant at .50 for all problems.

---

[8] H. D. Fishbein and P. Fingerman, The effects of perceptual organization on performance in classification tasks, *Psychon. Sci.*, 16, 1969, 76–77.

Procedure. At the outset each *S* was given detailed instructions describing the stimulus population, the nature of the concept to be learned, the method of responding, the meaning of the feedback, and the criterion of problem solution. The stimulus cards were shown to *S* one at a time. To each card, *S* responded 'yes' or 'no' to indicate his guess at the correct classification. Following each response, *E* provided feedback by telling *S* if he was right or wrong and by placing the card face up in the correct one of the two boxes labeled YES and NO. Thus *S* could normally observe the last card in each category. If the end of the stimulus deck was reached before *S* had met criterion, all cards were retrieved, the deck was reshuffled, and the task was resumed. The criterion of problem solution was 16 consecutive correct responses. Performance was self-paced in that *S* was given as much time as needed to respond, and instructions stressed accuracy rather than speed.

Results and Discussion

The results are summarized in Table II, which shows the mean number of errors for the six treatment combinations; the results for trials to last error were essentially the same. As is evident from the table, the type of solution had a marked effect on performance [$F(2, 66) = 7.55, p < .01$]; performance was best for the isolated-number solution, intermediate for the letters solution imbedded in an array of both numbers and letters, and worst for the same letters solution when only letters were present. A Newman-Keuls test of the means for solution types indicated that each mean differed significantly from the others (with, at most, $p < .07$). Despite the appearance of the data in Table II, neither the solution-spacing effect nor its interaction with solution type was significant.

TABLE II

MEAN NUMBER OF ERRORS FOR THE EXPERIMENTAL CONDITIONS: EXPERIMENT I

	Numbers	Letters embedded in numbers and letters	Letters embedded in letters
Adjacent	2.75	12.92	16.58
Spaced	7.42	12.00	22.83

The beneficial effect of isolating the relevant dimensions is consistent with earlier findings and provides still further confirmation of the von Restorff effect in concept learning. The improvement in performance from isolating the irrelevant dimensions is contrary to previous results, however, and contradicts a straightforward prediction from the theory of Trabasso and Bower (n. 1; see especially p. 148). The most appealing interpretation of this result is that offered at the outset, namely, that subjects are able to eliminate permanently the number dimensions because their uniqueness (saliency) facilitates memory for having tried them. There are, however, alternative explanations of this finding, and more detailed consideration is required.

First, because the method of isolation used here permitted subjects to partition the total set of dimensions into two subsets, it is possible that many subjects may have considered either numbers-only or letters-only solutions, but not 'mixed' letter-number solutions. If a subject, after trying the two numbers together, irrationally discards all solutions involving numbers, his set of 'tenable' hypotheses is artificially reduced; when there are no mixed solutions, as in this experiment, performance is likely to be improved relative to a nonisolated control group. Those subjects who second-guess the experimenter and start with letters gain a similar advantage; admittedly there cannot have been many such subjects or the von Restorff effect with numbers relevant would not have occurred. If this explanation is correct, a mixed one-number, one-letter solution should be learned more slowly than a two-letter solution when there are two isolated number dimensions present. In the case where there is only one number dimension, however, prediction is less clear unless it is assumed that subjects completely disregard the single number.

Among subjects who are willing to entertain mixed solutions at the outset, it is crucial whether or not they can recognize the relevance (or irrelevance) of number dimensions individually. For subjects who can detect the relevance of one number dimension while testing two, a mixed solution should be easier than a letters-only solution, since attention is directed to part of the solution immediately. Furthermore, mixed-solution performance would be better when there is only one number dimension present than when there are two, since the time wasted testing the other (irrelevant) dimension should generally be reduced. For subjects who are unable to recognize the relevance of one of the number dimensions, one might suppose that rejection of the two-number solution could be followed by elaborate testing of mixed-solution possibilities. However, relative to performance on letter-only arrays, such behavior would lead to poorer performance on letters solutions, an outcome specifically contradicted by the results of Experiment I.

The second possible explanation of the superiority of letters solutions when numbers were present involves the observation that the relevant letter dimensions were located adjacent to the irrelevant number dimensions. Since the subject's attention is drawn to that part of the array by the highly salient number dimensions, it is possible that they examined the letters in that part of the array, either incidentally while testing number solutions or deliberately when the number dimensions failed. This possibility is consistent with the results of experiments dealing with the von Restorff effect in serial learning, in which rate of learning items adjacent to the isolated item was increased relative to a list without the isolated item.[9] There is a direct test of this possibility; a two-letter solution located adjacent to two irrelevant number dimensions should be learned more quickly than one located in a different part of the array.

EXPERIMENT II

The purpose of this experiment was to test the various possible alternate interpretations of the results of Experiment I. Specifically, the following predictions were made. First, if subjects refuse to consider mixed number-letter solutions until late in the problem, then letters-only solutions should be easier than mixed solutions, regardless of whether one or two number dimensions are present. Second (assuming that most subjects examine the number dimensions first), if subjects can detect the fact that one of the number dimensions is relevant, performance with mixed-solutions should be superior to that with letters-only solutions; in addition, mixed solutions should be learned faster when there is only one number dimension than when there are two. Third, if the beneficial effect of irrelevant saliency in Experiment I was the result of directing the subjects' attention to the part of the array containing the relevant letter dimensions, then placing the irrelevant number dimensions in a separate part of the array should hurt performance.

---

[9] J. Wishner, T. E. Shipley, Jr., and M. S. Hurvich, The serial position curve as a function of organization, this JOURNAL, 70, 1957, 258–262.

Method

Subjects and design. The *S*s were 64 summer-session students at Kansas State University who were paid for participating. Sixteen *S*s were assigned to each of the four experimental treatment conditions in a stratified random order. The experimental design used was a simple randomized one-factor design with four levels of solution type. The problem solutions were (*a*) two letters, adjacent to two numbers in an eight-letter, two-number array, (*b*) two letters, well separated from two numbers in an eight-letter, two-number array, (*c*) one letter and one number, in an eight-letter, two-number array, and (*d*) one letter and one number, in a nine-letter, one-number array. Examples of positive instances of the four solutions can be seen in Table III, where the solution in each case is italicized.

TABLE III

EXAMPLES OF POSITIVE INSTANCES FOR EACH OF THE FOUR EXPERIMENTAL CONDITIONS: EXPERIMENT II

e	F	G	h	3	j	*5*	L	*m*	N
E	f	G	h	i	2	*K*	5	*m*	n
e	f	g	H	I	J	*5*	1	*m*	N
E	2	G	4	i	j	*K*	1	*m*	N

Task, materials, and procedure. The task, materials, and procedure for this experiment were the same as those of Experiment I, except that the stimuli contained ten dimensions. Specific examples of stimuli from each deck of cards are shown in Table III.

Results and Discussion

The results are summarized in Table IV, which shows mean number of errors for the four experimental treatment conditions. The results for trials data again were essentially the same. There was a marked difference in performance as a function of solution type [$F\ (3, 63) = 6.87, p < .01$]. Groups with mixed solutions performed significantly better than groups with two-letter solutions, a result confirmed by Newman-Keuls test ($p < .05$); however, the differences between the two mixed-solution groups, and between the two letters-only solution groups, were not significant.

Two of the experimental predictions were clearly contradicted by these outcomes. First, within the sensitivity of this experiment, performance with two-letter solutions was unaffected by placing the number dimensions closer or farther away in the array. Thus it does not appear that subjects necessarily follow their examination of the number dimensions by examining nearby letter dimensions.

TABLE IV

MEAN NUMBER OF ERRORS FOR THE EXPERIMENTAL CONDITIONS: EXPERIMENT II

Solution type			
One letter, one number, in eight-letter, two-number array	One letter, one number, in nine-letter, one-number array	Two letters adjacent to two numbers, in eight-letter, two-number array	Two letters not adjacent to two numbers, in eight-letter, two-number array
8.63	6.69	15.63	17.88

Second, the nearly equal performance of groups given mixed solutions, coupled with the finding that both of these groups performed better than groups given two-letter solutions, indicates that subjects do not restrict their hypothesizing to number-only and letter-only solutions. In fact, the uniform superiority of mixed solutions over letter-only solutions indicates that subjects are able to determine that one of the number dimensions is relevant, even though there may be others present which are irrelevant, and even though the remainder of the concept is yet to be discovered. This result is consistent with an earlier suggestion of Trabasso and Bower that relevant stimulus dimensions are learned separately in multidimensional conceptual problems.[10]

Taken together, the results of Experiments I and II indicate that subjects follow a fairly definite sequence or strategy in dealing with a conceptual problem of this type—one whose stimuli consist of an array containing a few highly salient dimensions (numbers) embedded in a larger number of less salient dimensions (letters). The first stage is the testing of numbers, leading to a von Restorff effect if numbers only are relevant. If, in this process, the subject discovers that one of the number dimensions is relevant, the next stage is the testing of mixed number-letter solutions involving that number dimension. Apparently, most subjects are able to detect a relevant number dimension even when the other is irrelevant. On the other hand, if both number dimensions are irrelevant, the subject appears to proceed directly to test letter combinations. Experiment I indicates that this stage is aided by the subject's ability to remember that the number dimensions are irrelevant.

---

[10] T. Trabasso and G. H. Bower, Component learning in the four-category concept problem, *J. math. Psychol.*, **1**, 1964, 143–169.

The data presented here suggest that previous formulations of the role of cue saliency must be modified to take into account the method by which increased saliency is obtained, for different methods may have quite different effects on both strategy and memory. At a minimum, there is evidence here that increasing the cue weight of irrelevant dimensions may have exactly the opposite effect from that which might be inferred from earlier work.

APPARENT SOUND DISPLACEMENT DURING VESTIBULAR STIMULATION

GENE LESTER, *Wheaton College*, and
RICARDO B. MORANT, *Brandeis University*

Abstract. Vestibular stimulation by means of angular acceleration around the midbody vertical axis produces a shift in the apparent position of a sound objectively rotating with the subject—the audiogyral illusion. Two possible explanations of the illusion were tested by a new method that leads to considerably reduced variance compared with previous methods. There were no significant differences in the extent of the illusion when stimulus tones of 500~, 1,024~, and 5,000~ were used. The illusion was significantly diminished in extent when the subjects strained their immobilized heads in the direction opposite from the usual apparent sound displacement during rotation. It is concluded that the illusion is based on a change in the felt position of the head rather than on differential hearing changes.

Vestibular stimulation induced by acceleratory or deceleratory rotation around the midbody vertical axis (Z axis) produces various illusory effects.[1] One of these, the audiogyral illusion, consists of the apparent displacement of a sound that is objectively stationary with respect to the subject.[2] The apparent displacement is opposite to the direction of rotation during acceleration and in the direction of original rotation during deceleration.[3]

Received for publication by Professor E. B. Newman. The investigation was supported by a National Science Foundation predoctoral grant to the first author and by Research Grant M-3658 from the National Institute of Mental Health, United States Public Health Service, to the second author. Dr. Lester is now at the State University College at Buffalo.

[1] I. P. Howard and W. B. Templeton, *Human Spatial Orientation*, 1966, 97–138.

[2] B. Clark and A. Graybiel, The effect of angular acceleration on sound localization: The audiogyral illusion, *J. Psychol.*, 28, 1949, 235–244.

[3] H. Munsterberg and A. H. Pierce, The localization of sound, in H. Munsterberg, Studies from the Harvard Psychological Laboratory, *Psychol. Rev.*, 1, 1894, 461–476; A. H. Pierce, *Studies in Auditory and Visual Space Perception*, 1901; B. Bourdon, Sensations causées par des rotations passives de tout le corps, *J. Psychol. (Paris)*, 30, 1933, 590–616; Clark and Graybiel (n. 2); M. D. Arnoult, Post-rotatory localization of sounds, this JOURNAL, 63, 1950, 229–236; L. B. W. Jongkees and R. A. van der Veer, On directional sound localization in unilateral deafness and its explanation, *Acta otolaryng.*, 49, 1958, 119–131.

Two hypotheses have been considered to explain the illusion. The first postulates that vestibular stimulation produces a difference in hearing between the two ears;[4] the second, that it produces a change in the spatial frame of reference.[5] Both can be stated in general form. The differential-hearing-change hypothesis suggests that vestibular stimulation produces a differential change in the sensitivity of the two ears, which results in apparent sound displacement. The frame-of-reference hypothesis suggests that vestibular stimulation produces a change in 'apparent straight-ahead' and that localization of the sound with respect to the *new* apparent straight-ahead results in apparent displacement.

A recent attempt to evaluate the two hypotheses by predicting the direction of the illusion when the source of sound is behind the subject or when one ear is occluded produced inconclusive results.[6] The present investigation thus had two purposes. The first was to establish a reliable method for measuring the illusion. The second was to use the method to evaluate further the two hypotheses proposed to account for the illusion.

A relevant point for the question of reliability of measurement is the fact that all previous quantitative studies of the audiogyral illusion have used abrupt deceleration from rotation to provide vestibular stimulation. Rapid deceleration from rotation produces violent stimulation of the vestibular apparatus and can have deleterious effects on vestibular thresholds.[7] Rapid deceleration probably also distracts attention from the sound-localization task. In view of this, the experiments reported here were conducted with gradual acceleration and deceleration. To further increase reliability, only subjects trained in sound localization within this specific experimental situation were used.

Implications drawn from each hypothesis were tested in two experiments. Experiment I examined two possibilities of the differential-hearing-change hypothesis: that apparent sound displacement is due to a change (*a*) in loudness threshold or (*b*) in latency of response. In either case, the sound should appear displaced in the direction of the more sensitive ear.[8] The implication is that the occurrence of differential hearing changes might be reflected in the amounts of apparent displacement reported for tones of different frequencies, since high-frequency tones are known to be localized on the basis of differences in intensity at the ears, low-frequency tones by comparisons of time or phase differences, and intermediate-frequency tones with reference to both kinds of cues.[9] If vestibular stimulation causes apparent interaural intensity differences, high-frequency tones might appear more displaced than low-frequency tones.[10] If vestibular stimulation causes apparent interaural time or phase differences, low-frequency tones might appear more displaced than high-frequency tones.[11] On the other hand, if the illusion is not caused by either kind of differential hearing change, equal displacements should be reported for all frequencies.

---

[4] O. Leisse, Sound localization and vestibular apparatus, *Arch. Ohr. Nas. Kehlkopfheik.*, 114, 1925–26, 118–128.

[5] Pierce (n. 3), 143; Jongkees and van der Veer (n. 3).

[6] Gene Lester and R. B. Morant, Sound localization during labyrinthian stimulation, *Proc. 75th annu. Conv. Amer. psychol. Assoc.*, 1967, 19–20.

[7] A. A. J. van Egmond, J. J. Groen, and L. B. W. van Jongkees, The turning test with small regulable stimuli: I, Method of examination; cupolometria, *J. Laryng.*, 62, 1948, 63–69.

Experiment II was designed to evaluate one possible form of the frame-of-reference hypothesis, that the audiogyral illusion is due to a change in the felt position of the head occasioned by the vestibular stimulation. The felt-head-position hypothesis is related to the explanation proposed by Pierce, who noted that the audiogyral illusion was related to the reflex tendency to rotate the head after the chair stopped turning.[12] He saw the audiogyral illusion as analogous to the paralyzed-eye-muscle illusion, in which, when the paralyzed eye is directed to the right but maintains its position with respect to some object, the object is perceived as having turned to the right.[13] In the case of the audiogyral illusion, the relevant innervation, according to Pierce, is of the neck muscles. He suggested that "the tension in the muscles of the neck is ordinarily the index of the head's position with respect to the body, and the distribution of sound intensities being of a particular sort, the sound must be misplaced in the same direction as the tendency to rotate the head. That is, the tension of the neck muscles without actual movement has been interpreted by assigning a greater angle to the auditory object. Of course there is no conscious process of inference here. It is all a matter of association and immediate perception. But the very fact that the tendency to move the head is not consciously known causes the illusion, for if it were known as such, a correction would be introduced" (n. 3, p. 143).

---

[8] W. E. Feddersen, T. T. Sandel, C. D. Teas, and L. A. Jeffress, Localization of high-frequency tones, *J. acoust. Soc. Amer.*, 29, 1957, 988–991; A. W. Mills, Lateralization of high-frequency tones, *J. acoust. Soc. Amer.*, 32, 1960, 132–134; G. W. Stewart, The function of intensity and phase in the binaural location of pure tones: I, Intensity, *Phys. Rev.*, 15, 1920, 425–431.

[9] S. S. Stevens and E. B. Newman, The localization of pure tones, *Proc. nat. Acad. Sci.*, 20, 1934, 593–596.

[10] Feddersen et al. (n. 8); Mills (n. 8).

[11] Stewart (n. 8).

[12] Pierce (n. 3), 143. Actually, Pierce does not distinguish acceleration from rotation. The adequate stimulus for vestibular functioning is angular acceleration or deceleration rather than rotation per se.

[13] E. Mach, *The Analysis of Sensation*, 1959, 128.

Pierce attempted to test his hypothesis by allowing the subject to turn his head in the direction of tension after rotation, then turn it back again and judge the location of the sound. This procedure, he assumed, would relieve the neck-muscle tension and satisfy the tendency to move. He rejected the hypothesis when he found that such voluntary head movements did not eliminate the illusion. Pierce went on to attribute the audiogyral illusion to eye nystagmus, but reversed the true directions of the slow and fast nystagmic components in his explanation.

Pierce's method for abolishing neck-muscle tension was probably not valid. In view of the lasting effects of abrupt deceleration, it seems unlikely that a single forward-and-back movement of the head satisfied the tendency to move (see n. 7). We must conclude that the hypothesis relating neck-muscle tension and the audiogyral illusion has not received a satisfactory test. Experiment II was directed toward testing the hypothesis; the tendency of the head to move in one direction was counteracted by a voluntary straining of the head toward the opposite direction.

EXPERIMENT I

Method

Subjects. Four undergraduate students trained in auditory localization served as *Ss.*

Apparatus. Angular acceleration was provided by means of a rotating chair. A wooden armchair was padded with foam rubber and mounted on a 1-in. rolled steel shaft. The shaft was turned by a ⅛ h.p. motor (Bodine NSH 54 RL) whose speed was regulated by a speed control (Minarik SH-52P). The chair could rotate in either direction. Acceleration by means of the speed control was found to be approximately linear when chair speed was timed with a stopwatch. The box containing the chair motor was bolted to a 3 × 3 in. sheet of 1-in. plywood, which in turn was bolted to the floor. A set of automobile seat belts restrained *S* in the chair. A U-shaped padded headrest and a biteplate holder were mounted on a shaft extending vertically from the back of the chair. Both the headrest and the biteplate were adjustable.[14] The *S*'s head was over the center of rotation and tilted slightly forward so that a line from the corner of the eye to the ear canal was horizontal. Thus, the horizontal canals were approximately in a horizontal position during testing.

Sound stimuli were provided by a Lafayette audio oscillator, a 1-in. speaker, and a battery-powered amplifier. The speaker, battery, and amplifier were mounted on a U-shaped wooden support clamped to the arms of the chair. The speaker was fixed 36 in. in front of *S*, at his waist level. Slip rings on the chair shaft connected the oscillator and speaker, so that *E* could turn the tones on and off at a control panel while chair was rotating. Three stimulus tones were used: 5,000~, 1,024~, and 500~.

The experimental room was fairly well insulated from external sounds by a double door. No precautions were taken to prevent echoing of sounds from the bare walls of the room, but, as is noted below, the *S*s were able to localize sounds during the control condition with high accuracy. During the experimental condition, *S* and the sound source were rotating within the room, so that *S* was constantly being exposed to new echoic conditions, with only the effect of the audiogyral illusion remaining constant. The acoustic conditions were such that when the experiments reported below were replicated with stimulus tones supplied through earphones, the results were very similar.

Procedure. Auditory localization training consisted of having *S* judge the position of a tone relative to the straight-ahead with his eyes closed, and then having him shown and told the sound's actual position. Tones at each of the three frequencies, positioned from 40 deg. left to 40 deg. right in 10-deg. intervals, were used. The frequency and position of the tones presented were randomized during two one-hour training sessions.

Following training, a series of 36 control readings (12 with each of the three tones) was taken with the sound presented straight ahead. These presentations were embedded in a series in which the same sounds were presented randomly in other positions. *S*s were instructed to be as accurate as possible, and although the sounds were objectively presented in 10-deg. steps, judgments of their position were frequently reported in 1-deg. steps. Accuracy of performance was not communicated to *S*. These readings were taken with *S* immobilized in the chair, and they served as the control readings for both Experiments I and II.

Each *S* in each experiment then underwent 72 rotations, half of them clockwise and the other half counterclockwise. The tone of each of the three frequencies was used 12 times with each direction of rotation. The order of testing was balanced for each *S*. Two judgments were made during each rotation, one during the accelerating phase and the other during the decelerating phase.

---

[14] Instead of biting onto the plate, *S* placed his upper lip against it so that the two prongs of the plate pressed on either side of the lip. This method of immobilizing the head was chosen in preference to the conventional biteplate procedure because *S*s in the preliminary studies complained of nausea aggravated by the presence of the biteplate in the mouth during rotation. (Similar difficulties were reported by W. C. Hixson, J. I. Niven, and C. A. Lowery, A head restraint device for vestibular studies, Document MR-005,04-0021,142, January 20, 1967, Naval Aerospace Medical Center, Pensacola, Florida.) Holding the head firmly between the headrest and the biteplate eliminated complaints of nausea and immobilized the head successfully.

The testing procedure was as follows. The sound was turned on and the chair accelerated from 0 to 20 rpm in 10 sec. (12 deg/sec^2). After a 1-sec. interval of steady rotation, the tone was turned off. *S* then judged the position the sound had occupied 1 sec. before being turned off. Rotation at 20 rpm was then continued until *S* reported that a 1-sec. pulsed sound appeared to come from straight in front. This judgment, reported after about 30 sec. of rotation, was used to determine when the effects of acceleration had passed. The test tone was then turned on, and the chair decelerated from 20 to 0 rpm in 10 sec. After a 1-sec. pause, the tone was turned off, and *S* judged the position it had occupied 1 sec. earlier. Before the next rotation, *S* sat still, with head immobile, until he reported that he no longer felt as if he were rotating. Three one-hour sessions on separate days were required to complete the experiment.

Results

Localization in the absence of vestibular stimulation: Control readings. In Table I the mean position in space that each tone was judged to occupy during the control readings is presented separately for each subject. The mean errors from objective straight-ahead were not significantly different from zero (t tests, $p > .05$), indicating that there was no lateral localization bias in the absence of vestibular stimulation. The average errors of localization were similar to those reported by Stevens and Newman for well-trained subjects (average error of 4.6 deg. for a 1,000~ tone; see n. 9).

Reduction of variability of the audiogyral illusion. It is difficult to compare the variability of the illusion found in the present investigation with that found by other workers, since average or standard deviations have not previously been reported. One measure which generally has been reported is the percentage of trials in which the illusion is not found. Pierce (n. 3, p. 141), with an

TABLE I

LOCALIZATION OF TONES AT THE MIDLINE DURING CONTROL READINGS: MEAN APPARENT DISPLACEMENTS, STANDARD DEVIATIONS, AND AVERAGE ERRORS

	5,000~			1,024~			500~		
Subject	$\bar{X}$	*SD*	*AE*	$\bar{X}$	*SD*	*AE*	$\bar{X}$	*SD*	*AE*
P. M.	0	11.5	8.3	.8 L	6.4	5.0	.8 R	4.7	2.5
F. F.	.8 L	8.6	4.2	3.4 R	7.4	5.0	.8 R	2.8	.8
J. M.	.8 R	2.8	.8	1.7 R	5.5	3.5	1.7 L	5.5	3.5
D. C.	0	4.3	1.7	.8 L	4.9	2.5	2.5 R	4.9	2.5

Note: All measures in degrees. R (right) and L (left) show direction of apparent displacement.

unspecified sound, found no displacements on approximately 6% of the trials and displacements opposite to the usual direction of the illusion on 21% of the trials. Clark and Graybiel (n. 2), using a 1,024~ tone and other conditions very similar to the present experiment, but with deceleration only, reported no displacements on 20% of the trials and opposite displacements on 9%. Arnoult (n. 3), with a 500~ tone, reported no displacement on 15% of the trials and opposite displacement on 9%. In contrast, the present experiment found, for the 576 trials run with the three tones, no displacement on less than 6% of the trials and opposite displacement only twice.

Comparison of displacement effects for the three frequencies. The means and standard deviations of the apparent sound displacements reported by each subject for each tone are presented in Table II. The displacement for each experimental condition is in the direction reported by previous workers. During clockwise acceleration and deceleration from counterclockwise rotation, the sounds were displaced left. During counterclockwise acceleration and deceleration from clockwise rotation, the sounds were displaced right. These means were significantly different from the control readings beyond the .01 confidence level (*t* tests). Computation of the overall mean displacements irrespective of direction of labyrinthian stimulation yielded the values 500~, 15.0 deg.; 1,024~, 15.0 deg.; 5,000~, 16.0 deg.

The results of a three-way analysis of variance testing the effects of direction of rotation (clockwise or counterclockwise), stage of rotation (acceleration or deceleration), and frequency of the sound stimulus are summarized in Table III. The variances were homogeneous as determined by Bartlett's test.[15] The only significant source of variance in Table III is the interaction between direction and stage of rotation, which represents the basic audiogyral illusion. The lack of significance of the main effects indicates that equivalent illusory effects were obtained with clockwise and counterclockwise rotation, and with acceleration and deceleration.[16] Of

---

[15] A. L. Edwards, *Experimental Design in Psychological Research,* 1962.

[16] Since apparent displacements of the sound to the right were assigned a + value and displacements to the left were assigned a — value, a low *F* ratio indicates that displacements in each direction summate to cancel each other out. As indicated above, all mean displacements during vestibular stimulation were significantly different from control. Because of the +/— value convention, high *F* ratios occur only when + and — values are summed separately in the interaction terms.

TABLE II

LOCALIZATION OF TONES AT THE MIDLINE UNDER VESTIBULAR STIMULATION DURING EXPERIMENT I: MEAN APPARENT DISPLACEMENTS AND STANDARD DEVIATIONS

	5,000~				1,024~				500~			
	Clockwise		Counterclockwise		Clockwise		Counterclockwise		Clockwise		Counterclockwise	
Subject	Accel.	Decel.	Accel.	Decel.	Accel.	Decel.	Accel.	Decel.	Accel.	Decel.	Accel.	Decel.
P. M.												
$\bar{X}$	12.9 L	12.1 R	12.5 R	9.0 L	12.9 L	11.4 R	11.3 R	10.0 L	11.8 L	10.4 R	13.3 R	11.2 L
SD	3.2	3.7	3.2	3.7	4.3	4.7	2.0	3.5	2.5	3.5	3.2	4.7
F. F												
$\bar{X}$	22.5 L	27.5 R	27.1 R	24.5 L	17.1 L	30.0 R	29.2 R	25.4 L	17.1 L	29.6 R	29.6 R	24.2 L
SD	6.3	7.2	8.2	8.5	9.9	5.4	9.7	6.6	10.4	4.6	8.4	10.3
J. M												
$\bar{X}$	5.8 L	15.8 R	30.0 R	13.1 L	8.9 L	10.7 R	10.4 R	10.0 L	7.9 L	11.7 R	9.2 R	8.8 L
SD	6.4	8.7	8.1	5.9	4.8	6.5	7.2	6.8	6.3	7.9	4.4	2.8
D. C.												
$\bar{X}$	10.8 L	11.7 R	6.7 R	14.2 L	12.1 L	15.4 R	10.4 R	14.2 L	10.4 L	17.1 R	10.0 R	18.8 L
SD	3.5	8.3	4.6	5.9	3.7	2.5	1.5	4.8	3.2	7.7	2.0	6.0

Note: All measures in degrees. R (right) and L (left) show direction of apparent displacement.

TABLE III

ANALYSIS OF VARIANCE: EXPERIMENT I

Source	*df*	*MS*	*F*
Direction (A)	1	22.8	
Stage (B)	1	13.8	
Frequency (C)	2	4.8	
A × B	1	11,190.5	179.9*
B × C	2	2.2	
A × C	2	11.8	
A × B × C	2	3.2	
Within cells	36	62.2	

* $p < .01$.

more immediate theoretical importance is the finding that an equivalent audiogyral illusion was obtained for high-, low-, and intermediate-frequency tones—a finding that supports neither mechanism suggested by the differential-hearing-change hypothesis.

EXPERIMENT II

The pattern of results obtained in Experiment I is congruent with Pierce's explanation that the audiogyral illusion is due to a reflex tendency to rotate the head during vestibular stimulation. In Experiment II an attempt was made to counteract the effects of this tendency by having the subject strain his head in the direction opposite to the reflex. It was hypothesized that counteracting the tendency to move the head would reduce the illusion and that the reduction would be equivalent for all three tones.

Method

Each *S* was instructed to strain his head toward the right during clockwise acceleration and deceleration from counterclockwise rotation, and toward the left during counterclockwise acceleration and deceleration from clockwise rotation. These directions of strain counteract the direction of reflex head movement during the various conditions of vestibular stimulation.[17] 'Head strain' consisted of having *S* try to rotate his immobilized head toward one side when the stimulus tone was on. He was to relax when the stimulus tone went off and then to judge the tone's terminal position. The headrest and biteplate held the head firmly, so that actual rotation of the head did not occur.

---

[17] E. Buys, Contribution à l'étude du nystagmus oculaire de la rotation chez l'homme, *Rev. d'oto- neuro-ocul.*, 2, 1924–25, 641–659, 721–749; 3, 1924–25, 10–32, 105–126; A. Landau, Ueber einen tonischen Lagereflex beim älteren Saügling, *Klin. Wschr.*, 2, 1923, 1253–1255.

With the exception of having *S* strain his head during vestibular stimulation, the conditions of stimulation, the procedures, and the design were identical to those of Experiment I. The same four students served as *Ss.* A period of at least one day intervened between the last session of Experiment I and the first session of Experiment II.

Results

The means and standard deviations of sound displacement reported by each subject for each tone are presented in Table IV. Comparison of these results with those of Experiment I indicates that the magnitude of the audiogyral illusion was substantially reduced with the introduction of contrary head strain; the data showing the magnitude by which the illusion was reduced over all experimental conditions for each *S* is presented in Table V.

When the results of Experiments I and II were combined, the variances were homogeneous, as determined by Bartlett's test (n. 15). A four-way analysis of variance was used to analyze the combined results of Experiments I and II, examining the effects of experimental treatments (presence or absence of head strain), direction of rotation, stage of rotation, and frequency of stimulus tone. The F ratios for the main effects were not statistically significant. This indicates that in both experiments (with and without head strain during rotation), displacements to the right were equivalent in size to displacements to the left (see n. 16). Statistical evaluation of the effect of head strain is derived when direction and stage of rotation are held constant, that is, in the interaction of treatment, direction, and stage. This interaction term was highly significant ($F = 61.6$, df 1, 95): the apparent displacements reported in Experiment II (with head strain) were significantly smaller than those reported in Experiment I (without head strain). All other F ratios were nonsignificant, except the interaction of direction and stage, which represents the basic audiogyral illusion and which was highly significant ($F = 222.4$, df 1, 95). We conclude that when a change in apparent head position counteracts the normal apparent displacement of the head during vestibular stimulation, the extent of the audiogyral illusion is significantly diminished.

Since all the judgments for Experiment II were made after Experiment I was performed, the question of a simple order or habituation effect had to be raised. Analysis of the data within each

TABLE IV

LOCALIZATION OF TONES AT THE MIDLINE UNDER VESTIBULAR STIMULATION AND CONTRARY HEAD STRAIN DURING EXPERIMENT II: MEAN APPARENT DISPLACEMENTS AND STANDARD DEVIATIONS

	5,000~				1,024~				500~			
	Clockwise		Counterclockwise		Clockwise		Counterclockwise		Clockwise		Counterclockwise	
Subject	Accel.	Decel.	Accel.	Decel.	Accel.	Decel.	Accel.	Decel.	Accel.	Decel.	Accel.	Decel.
P. M.												
$\bar{X}$	.2 L	.4 R	.4 L	0	.4 L	.4 R	.2 R	1.2 L	.2 L	.4 R	.4 R	.3 L
SD	.5	.9	1.0	1.2	1.0	1.3	.8	4.2	.5	1.3	1.3	.8
F. F.												
$\bar{X}$	4.2 L	7.9 R	7.1 R	7.9 L	3.7 L	10.4 R	5.0 R	1.6 L	1.6 L	6.6 R	5.8 R	2.1 L
SD	3.4	6.3	7.8	8.8	5.1	7.2	6.4	3.7	2.4	3.8	7.1	4.7
J. M.												
$\bar{X}$	1.1 R	3.6 R	5.5 R	.2 L	1.8 L	2.8 R	1.3 R	1.9 L	1.1 L	2.4 R	1.1 R	2.4 L
SD	3.1	3.6	2.9	2.0	2.3	3.0	1.7	2.3	1.8	2.0	1.8	2.0
D. C.												
$\bar{X}$	8.3 L	6.7 R	7.5 R	7.5 L	11.7 L	12.5 R	10.0 R	13.3 L	9.2 L	11.2 R	13.3 R	12.5 L
SD	6.6	5.5	4.7	5.2	3.5	3.2	3.5	5.6	2.7	4.2	4.8	5.6

Note: All measures in degrees. R (right) and L (left) show direction of apparent displacement.

TABLE V

MEAN APPARENT DISPLACEMENTS UNDER VESTIBULAR STIMULATION WITH AND WITHOUT CONTRARY HEAD STRAIN

Subject	Without head strain: Experiment I	With head strain: Experiment II
P. M.	11.56	.37
F. F.	25.31	5.32
J. M.	11.85	1.91
D. C.	12.65	10.30

Note: All measures in degrees.

experiment indicated no habituation effect over the successive days of testing.

DISCUSSION

Experiment I showed that the magnitude of the audiogyral illusion was unaffected by sound frequency. This result does not support the differential-hearing-change hypothesis but is congruent with the frame-of-reference hypothesis. Experiment II tested one specific formulation of the frame-of-reference hypothesis, the formulation which suggests that neck-muscle tension accounts for the audiogyral illusion. This result demonstrated that voluntary head strain significantly influenced the magnitude of the audiogyral illusion. We propose that both sets of results were due to changes in the felt position of the head induced by vestibular stimulation and voluntary head strain. There are two parts to the argument. The first has to do with the relationship between the felt position of the head and sound localization; the second, with the relationship between head strain and the felt position of the head.

In the first place, it is evident that egocentric sound localization must depend not only on differential sound cues at the ears but also on information regarding the position of the head relative to the body. If the head feels as though it has turned from the body midline, then sound is localized with respect to the new apparent head position. A sound directly in front of the head appears turned from the midline by whatever angle the head feels turned.

In the second place, lateral head strain induced either reflexly, through vestibular stimulation, or voluntarily, by turning the head against an immovable restraint, induces a change in the felt position of the head. In both instances the head feels turned in the direction of strain even though it is objectively straight ahead. The results of Experiment II are interpreted as a demonstration that reflexly induced neck-muscle tension can be counteracted by voluntarily straining the head in the opposite direction. We propose that the counteraction of the reflex tendency to move reduces the extent to which the head feels turned and consequently reduces the magnitude of the audiogyral illusion.

A TEST OF A PERSPECTIVE THEORY OF GEOMETRICAL ILLUSIONS

Dominic W. Massaro and Norman H. Anderson,
University of California, San Diego

Abstract. Three-dimensional Müller-Lyer figures were used in three experiments to test a perspective theory of geometrical illusions. The real depth cues were in one of two orientations, designed either to support or to oppose the action of the perspective depth cues hypothesized by the perspective theory. If any illusion is obtained with these three-dimensional figures, the theory implies, its magnitude should be different in the two orientations. In fact, a substantial illusion was obtained, but the difference between the two orientations was small and opposite to the prediction of perspective theory.

Perspective explanations of geometrical illusions have a long history,[1] but only recently has this approach received the conceptual clarification needed to open it to experimental analysis. Gregory's formulation of perspective theory rests on two main assumptions.[2] First, geometrical-illusion figures such as the Müller-Lyer ones in Figure 1ab are considered two-dimensional projections of three-dimensional figures; thus, they contain perspective cues for depth. Second, these cues are assumed to trigger a constancy-scaling mechanism that corrects for decrease in the retinal image with increasing distance. Those parts of the figure that would be more distant in the three-dimensional representation are enlarged and those that would be closer are diminished, in accord with size constancy.

This version of perspective theory has been applied especially to

---

Received for publication April 27, 1970. The first author, who is now at the University of Wisconsin, was supported by a postdoctoral fellowship, MH-39369-02, from the National Institute of Mental Health. Support was also provided by a National Science Foundation grant, GB-6666, to the second author. The authors thank Susan Keeve for her help.

[1] R. S. Woodworth, *Experimental Psychology,* 1938, 644–645; R. L. Gregory, *Eye and Brain,* 1966, 141–156.

[2] R. L. Gregory, Distortion of visual space as inappropriate constancy scaling, *Nature,* 199, 1963, 678–680; R. L. Gregory. Seeing in depth, *Nature,* 207, 1965, 16–19.

the planar Müller-Lyer figures in Figure 1ab.[3] As Gregory notes, these can be seen as two-dimensional projections of different objects depending on the direction of the wings. Figure 1a could be the projection of the near vertical corner of a building or of a box. The wings represent the lines of the top and bottom of the object receding into the distance. Size constancy, corresponding to the adjustment in apparent size that keeps the object parallelepipedal, would decrease the length of the main axis of the figure. In Figure 1b, which could represent the projection of the far corner of a room, the constancy scaling would analogously expand the apparent size of the main axis.

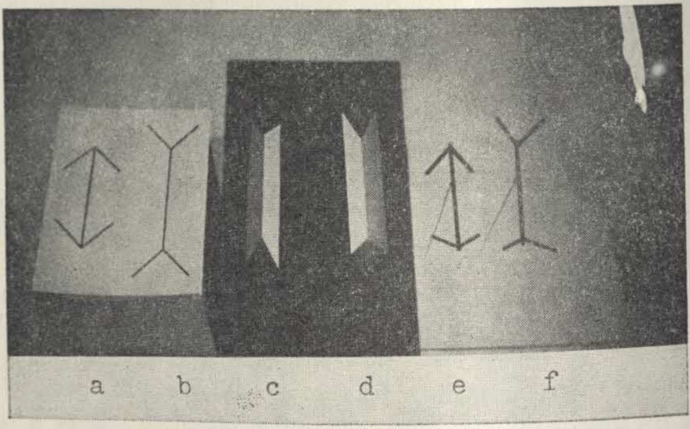


FIG. 1. (a) and (b) Typical two-dimensional Müller-Lyer figures; (c) three-dimensional figure with the wings oriented away from the observer, (d) wings oriented toward the observer; (e) stick figure with the wings oriented away from the observer, and (f) wings oriented toward the observer

Gregory's formulation has provoked numerous criticisms, most of which attempt to show that perspective theory does not account for certain illusions, such as the dumbbell version of the Müller-Lyer figure.[4] However, perspective cues could still be operative in some illusions, even though additional factors might also be involved. Perspective theory may thus have a valid role in the explanation of illusions, and the present experiments attempted to get a direct test of it by varying the depth cues of three-dimensional Müller-Lyer figures as is illustrated in Figure 1. The question of interest is how the real depth of the wings affects the normal illusion. For example, one critical comparison is between Figure 1c and 1d. These are geometrically identical, except that the wings extend away from the observer in Figure 1c, toward the observer in Figure 1d.

---

[3] Gregory, Seeing in depth (n. 2).

[4] R. H. Day, Inappropriate constancy explanation of spatial distortions, *Nature,* 207, 1965, 891–893.

Oddly enough, the most direct interpretation of perspective theory predicts no illusion for three-dimensional Müller-Lyer figures. Since they are easily seen in correct depth perspective, the hypothesized constancy-scaling mechanism should not be triggered inappropriately, and hence there should be no illusion. But this is wrong, as can perhaps be seen by comparing Figure 1e and 1f, as demonstrated below, and as reported earlier by Lucas and Fisher for a three-dimensional Poggendorff illusion.[5] It is necessary, therefore, to search for the perspective cues that must exist if perspective theory is to account for the illusion in the three-dimensional figures.

One possible application of perspective theory would attribute the illusion to inappropriate constancy scaling triggered by the three-dimensional form. Thus, Figure 1c represents the corner of a building no less than Figure 1a. Similarly, Figure 1d represents the corner of a room no less than Figure 1b. Direct transposition of the two-dimensional theory would then predict that the central axis should appear shorter in Figure 1c than in Figure 1d. Indeed, this interpretation would require that Figure 1c appear shorter than even the control version of Figure 1d with 90-deg. wings. As Figure 1 suggests, and as will be seen below, this is not true.

There seems to be only one remaining possibility, namely, that the illusion is produced by certain depth cues of the two-dimensional projections of the figures. The retinal image itself is two-dimensional, so the correct perception, or real depth, must be constructed from the available depth cues. These include the usual cues to real depth, such as convergence, stereopsis, interposition, and so forth. And, if perspective theory is correct, they also include the perspective features of the two-dimensional retinal image that are assumed to produce constancy scaling.

There are, then, two sets of depth cues, one 'real' and one hypothesized by perspective theory. These two sets of cues can act in concert or in opposition. If the real depth cues oppose the perspective cues, then the illusion should be less than if they act in concert. Accordingly, the figures shown in Figure 1cd provide a critical test. In Figure 1c, the wings slope away from the observer, and, therefore, the depth cues of the two dimensional projections disagree with the actual depth; in Figure 1d, they agree. Since the perception of depth is dependent upon all available cues, perspective theory predicts a larger illusion for Figure 1d than 1c. But if perspective theory is wrong, both figures should give essentially the same effect. It is on this logic that the experiments were designed.

---

[5] A. Lucas and G. H. Fisher, Illusions in concrete situations: II, Experimental studies of the Poggendorff illusion, *Ergonomics*, 12, 1969, 395–402.

METHOD

Stimuli for Experiments I and II. Examples of the stimuli are shown in Figure 1cd. They were constructed from stiff off-white cardboard, first cut to an areal version of the Müller-Lyer figure, and then folded along the central axis so that the two sides met at right angles. The width of the side was 32 mm. in all figures.

Twenty figures were used, in a 2 × 2 × 5 design. There were two orientations, away from or toward the *S*, as in Figures 1c and 1d respectively. Two lengths of central axis were used, 127 and 152 mm. And the edge of the side, or wing, met the central axis at one of five angles: 45, 60, 90, 120, and 135 deg. The angles shown in Figure 1cd are both at 135 deg.

These figures were mounted in wooden holders 25 cm. long, with the 5 × 3 cm. cross section shaped at one end to a 90-deg. angle to accommodate the figures. These holders rested on a table at eye level about 100 cm. from the *S*, with the figures projecting 10 cm. beyond the edge of the table. Two lamps were used in addition to the overhead fluorescent to provide uniform illumination on the figures. The *E* sat at the side of the table, presenting the stimuli and recording the responses. The *S*s were instructed to judge the apparent length of the central axis, and two practice trials were given.

Procedures for Experiments I and II. In Experiment I, the figures were presented in a horizontal position. The *S* made his response by adjusting the length of a motorized white tape displayed horizontally on a black panel about 30 cm. below and 10 cm. to the side of the stimulus figure. Two groups of 8 *S*s from introductory courses were used. Each *S* judged the 10 figures with one length of central axis. These 10 figures were judged 10 times in successive blocks, each block given in a different random order for each *S*.

In Experiment II, the figures were presented in a vertical position, as in Figure 1. The 16 *S*s judged all 20 figures, 5 times each in separately randomized order in five successive blocks. The response was made by marking off on a 205-mm. black vertical line beginning 11 mm. below the top edge of a sheet of white paper.

Stimuli for Experiment III. This experiment used black stick figures, 150 mm. long and 3 mm. thick, with 50-mm. wings. The wings were displaced 3 mm. in from the ends of the central axis to avoid any possible doubt about the location of the ends of the central axis. The figures were presented vertically on a knitting needle 7 cm. from a white background wall, about 110 cm. from the *S*.

The 5 three-dimensional stimuli were constructed so that they could fit in the corner of a room, the central axis in the vertical and the wings flat on the two walls. The wings had angles of 30, 60, 90, 120, and 150 deg. with the central axis. Figure 1ef shows the 60-deg. angle oriented away from, and the 120-deg. angle oriented toward, the camera. Two two-dimensional stick figures, analogous to those in Figure 1ab, and a control without wings, were also used.

Procedure for Experiment III. The general procedure was similar to that of Experiment II, and *S*s made their response by marking a line on a response sheet in the same way. The 5 three-dimensional figures were presented in both orientations, so there were 13 possible stimuli altogether. These were presented 5 times each in successive blocks. The *S*s were 20 students from an introductory course.

Statistical analysis. Complete three-way or four-way analysis of variance was applied to the data of each experiment, including trial blocks as a factor. The main results are obvious in the graphs, and only those that require special mention are specifically discussed.

RESULTS

Experiments I and II. The mean judgments in Experiments I and II are shown in the left and right panels of Figure 2. The upward trend of the curves shows that the apparent length of the constant central axis was markedly influenced by wing angle. Thus, the Müller-Lyer illusion was obtained with these three-dimensional figures.

Since the illusion did occur, the critical test of perspective theory is given by the comparison between the open and closed circles in each panel. If perspective theory is correct, the open circles, which correspond to the corner of a building or a box, should lie below the closed circles. Exactly the opposite was seen in both experiments. The mean difference between the inward and outward orientations was .64 mm. in Experiment I, and 1.64 mm. in Experiment II. The former difference was not significant [F (1, 14) = 3.22], but the latter was [F (1, 15) = 21.71]. Both experiments, therefore, contradict perspective theory.

Experiment III. Although the first two experiments argue against perspective theory, no cause was seen for the opposite effect obtained in Experiment II. The slight difference in retinal angle of the total figure would actually favor perspective theory. This was small, but it was feared that it might reflect some peculiarity of the cardboard figures. If such an effect existed, it might mask a

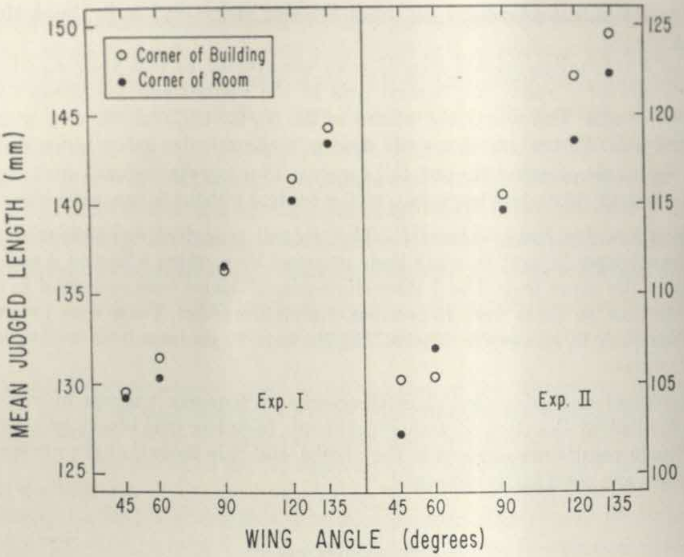


FIG. 2. Mean judged length as a function of the wing angle and the real depth perspective of the figures used in Experiments I and II

real perspective effect. Accordingly, stick figures were used to eliminate any possible perceptual ambiguity about the endpoints of the central axis.

Figure 3 shows the mean judgments of the three-dimensional stick figures. The upward slope of the curve demonstrates that these stick figures also produced the illusion. In the critical comparison, however, the mean difference between the inward and outward orientation was only .17 mm., which did not approach significance. Thus, these data also contradict perspective theory.

Miscellaneous results. A few minor results from the analyses should also be reported. In Experiment I, there was a significant increase over trials in mean response, from 132.5 mm. in the first trial block to 139.6 mm. in the last trial block. This increase may have been caused by the response mechanism, which was somewhat inconvenient to control. At any rate, no such increase was obtained in the other two experiments, in which the line-marking response was employed.

In Experiment III, the two-dimensional stick figures also showed the illusion, the mean judgments being 128.7 and 135.7 mm. for the inward and outward wings. In comparison, the mean judgment for the wingless control stick was 130.8 mm. These data show an asymmetry similar to the nonlinearity in Figure 3. Possibly, this reflects a contrast effect, as mentioned by Müller-Lyer.[6]

Experiment I showed a significant decrement in the magnitude of the illusion over trials. This result agrees with the practice decrement reported for two-dimensional figures.[7] The illusion still persisted at the last trial block, however, and no decrement was found in Experiment II. Experiment III showed a small, but nonsignificant, decrement in the illusion for both the two- and three-dimensional stick figures.

DISCUSSION

The present three-dimensional Müller-Lyer figures appear to provide a critical test of Gregory's perspective theory of geometrical illusions (n. 2). If this theory is to account for the illusion obtained with three-dimensional figures, it presumably must invoke the same set of depth cues assumed to operate in the two-dimensional figures.

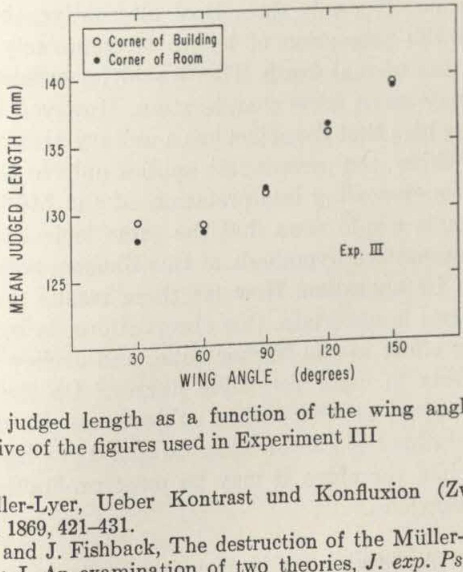


Fig. 3. Mean judged length as a function of the wing angle and the real depth perspective of the figures used in Experiment III

---

[6] F. C. Müller-Lyer, Ueber Kontrast und Konfluxion (Zweiter Artikel), *Z. Psychol.*, 10, 1869, 421–431.

[7] W. Köhler and J. Fishback, The destruction of the Müller-Lyer illusion in repeated trials: I, An examination of two theories, *J. exp. Psychol.*, 40, 1950, 267–281; R. H. Day, The effects of repeated trials and prolonged fixation of error in the Müller-Lyer figure, *Psychol. Monogr.*, 76, 1962, No. 14 (Whole No. 533).

These cues are hypothesized to produce a constancy correction for implied distance that expands or contracts the apparent length of the central axis.

But three-dimensional figures also contain a second set of depth cues, cues that produce the perception of their real depth. The critical comparison is, then, between the same figure in its outward and inward orientations, as illustrated in Figure 1c and 1d respectively. In the outward orientation, the two sets of depth cues disagree; in the inward orientation, they agree. Perspective theory then implies that the size of the illusion should be less in the former case than in the latter. It should be noted that this is a direct test of the perspective/constancy-scaling hypothesis. Other factors might play a role, but if perspective has any effect at all, then it should have appeared in the relevant comparison. Contrary to perspective theory, no such effect was obtained.

Perspective theory can hardly argue that the hypothesized constancy scaling does not apply to the three-dimensional figure, for then the theory cannot account for the strong illusion that these figures give. A new explanatory principle would then be required, one that presumably would explain the two-dimensional illusion at the same time. The only theoretical alternative, then, would be to argue that the perception of length is completely independent of the perception of real depth. This would be sufficiently remarkable that it may merit some consideration. However, it is not consistent with the idea that the object has a unitary perception.

Strictly speaking, the present test applies only to Gregory's perspective/constancy-scaling interpretation of the Müller-Lyer illusion. However, it would seem that the same logic would apply to the general perspective hypothesis of this illusion, regardless of the mechanism of its operation. How far these results may generalize to other illusions is uncertain. Our observations do indicate a constancy-scaling effect in the Necker cube, and Jeffrey has reported analogous effects in other reversible figures.[8] On the other hand, Fisher has reported no effect when subjects assume different perspective orientations of the ordinary Müller-Lyer figure.[9] For the standard illusion, therefore, it may be most profitable to seek the explanation elsewhere.

---

[8] H. J. Jeffrey, An inquiry into inappropriate constancy scaling, *Quart. J. exp. Psychol.*, 20, 1968, 294–296.

[9] G. H. Fisher, An experimental and theoretical appraisal of the inappropriate size-depth theories of illusions, *Brit. J. Psychol.*, 59, 1968, 373–383.

There is another approach to the illusions that has nearly as long a history as perspective theory. These are the 'total impression' theories that emphasize judgmental processes rather than sensory-perceptual factors.[10] Gregory has criticized the total-impression theories as being nontheories, incapable of definite prediction.[11] More recently, however, two mathematical formulations of this type have been suggested. Restle and Merryman have applied adaptation-level theory to the analysis of certain illusions.[12] Anderson has given a development based on integration theory and functional measurement.[13] Both these formulations make quantitative predictions that have enjoyed some provisional success.

---

[10] Woodworth (n. 1), 645.

[11] Gregory (n. 1), 142–143.

[12] F. Restle and C. T. Merryman, An adaptation-level theory account of a relative-size illusion, *Psychon. Sci.*, 12, 1968, 229–230.

[13] N. H. Anderson, Averaging model applied to the size-weight illusion, *Percept. & Psychophys.*, 1970 (in press).

A VISUAL SPATIAL AFTEREFFECT OF SURFACE SLANT

PETER M. WENDEROTH, *University of Sydney, Australia*

Abstract. The long (binocular) inducing period required by Bergman and Gibson to generate a measurable slant aftereffect was not necessary for a high-contrast surface. The aftereffect for the black-on-white lattice pattern dissipated exponentially. The complex surface, however, produced concomitant interfering effects, possibly decreasing the precision of aftereffect measures. A simpler surface, of vertical contours, sufficed for relatively accurate binocular (but not monocular) adjustments to the vertical. This pattern also produced interfering effects, possibly accounting for the correlation of aftereffect variance with inducing-figure slant. The five studies suggest using a single vertical contour as both inducing and test figures to study a hypothesized relationship between tilt aftereffect and slant aftereffect.

A sensory spatial aftereffect occurs when the judgment of some spatial property of a test stimulus is modified as a result of prolonged stimulation by an inducing stimulus. The aftereffect is assumed to be a simple additive component and is defined as the algebraic difference between a prestimulation judgment (pretest) and a poststimulation judgment of the test stimulus (posttest).[1]

Spatial aftereffects in the third dimension of visual space were first demonstrated by Köhler and Emery.[2] They reported, for example, that a frontal-parallel cardboard strip was judged to be slanted in one direction following fixation of a similar strip slanted in the opposite direction. Later, Bergman and Gibson reported the occurrence of a slant aftereffect induced by an irregularly textured burlap surface with its edges obscured from view.[3] The mean after-

Received for publication by Professor E. B. Newman, August, 1967. The research was part of a thesis awarded the M.A. at the University of Sydney, 1968. The author wishes to thank Professor R. H. Day for his assistance throughout the project.

[1] J. J. Gibson, Adaptation after-effect and contrast in the perception of curved lines, *J. exp. Psychol.*, 16, 1933, 1–31; W. Köhler and H. Wallach, Figural after-effects: An investigation of visual processes, *Proc. Amer. phil. Soc.*, 96, 1944, 269–357. Later investigations have usually not departed from this convention.

[2] W. Köhler and D. A. Emery, Figural after-effects in the third dimension of visual space, this JOURNAL, 60, 1947, 159–201.

[3] R. Bergman and J. J. Gibson, The negative after-effect of a surface slanted in the third dimension, this JOURNAL, 72, 1959, 364–374.

effects induced by slants of 15, 30, and 45 deg. were 2.65, 2.85, and 2.75 deg. respectively, with variances of 1.81, 4.03, and 2.11.[4] The purpose of the studies reported below was to investigate more extensively aftereffects of the type reported by Bergman and Gibson.

Gibson initially outlined a principle of 'normalization' to account for two-dimensional 'negative' aftereffects.[5] Köhler claimed that both the effects described by Gibson and the 'figural' aftereffects reported by Köhler and Wallach could be explained by a cortical-satiation hypothesis.[6] Köhler and Emery extended this claim to aftereffects in the third dimension (n. 2). Bergman and Gibson argued that their slant aftereffect could not be explained by the satiation hypothesis, since the effect was induced by texture gradients without visible contours (n. 3; see p. 367). However, where there is discriminable texture there must be discriminable contour.[7] Hence, this argument for the use of an irregularly textured, low-contrast surface is irrelevant.

There are at least two reasons for preferring to use a high-contrast, regularly textured surface. First, it would be expected that a measurable aftereffect could be induced with a stimulation period much less than the 4-min. minimum which Bergman and Gibson reported (n. 3; see p. 368). Second, regular surfaces allow more accurate judgments of slant.[8] Both factors would be expected to reduce aftereffect variance, the first by reducing the subject's fatigue during and after stimulation, and the second by resulting in less variable adjustments to the vertical. These considerations are important because the stimulus surface used by Bergman and Gibson induced small effects with relatively large variances. Hence, their tests of mean differences lacked power.[9]

---

[4] Bergman and Gibson reported the sum of a pair of aftereffect measures for each subject. The above recalculations are based on means.

[5] J. J. Gibson, Adaptation with negative after-effect, *Psychol. Rev.*, 44, 1937, 222–244; J. J. Gibson and Minnie Radner, Adaptation, after-effect, and contrast in the perception of tilted lines: I, Quantitative studies, *J. exp. Psychol.*, 20, 1937, 453–467; J. J. Gibson, Adaptation, after-effect, and contrast in the perception of tilted lines: II, Simultaneous contrast and areal restriction of the after-effect, *J. exp. Psychol.*, 20, 1937, 552–569.

[6] W. Köhler, Unsolved problems in the field of figural after-effects, *Psychol. Rec.*, 15, 1965, 63–83; Köhler and Wallach (n. 1).

[7] P. M. Wenderoth, The negative after-effect in the third dimension and the distance effect, unpublished B.A. thesis, University of Sydney, 1963. I. P. Howard and W. B. Templeton, The effect of steady fixation on the judgment of relative depth, *Quarterly J. exp. Psychol.*, 16, 1964, 193–203.

[8] J. J. Gibson, The perception of visual surfaces, this JOURNAL, 63, 1950, 367–384; W. C. Clark, A. H. Smith, and A. Rabe, The interaction of surface texture, outline of gradient, and ground in the perception of slant, *Canad. J. Psychol.*, 10, 1956, 77–81; H. R. Flock and A. Moscatelli, Variables of surface texture and accuracy of space perceptions, *Percept. mot. Skills*, 19, 1964, 327–334; A. L. Kraft and W. A. Winnick, The effect of pattern and texture gradient on slant and shape judgments, *Percept. & Psychophys.*, 2, 1967, 141–148.

EXPERIMENT I

The aim of Experiment I was to establish whether a measurable slant aftereffect occurred following a 2-min. fixation of a black-on-white lattice pattern.

Method

Apparatus. The apparatus was essentially the same as that used by Bergman and Gibson (n. 3), except for the stimulus surface. A black lattice pattern was drawn with india ink on a sheet of white pasteboard. All lines were 4 mm. wide with 20-mm. interspaces. A horizontal white fixation line, 4 mm. long and 1 mm. wide, was drawn at the intersection of a vertical and horizontal contour at the axis of rotation.

The *S* viewed the surface through circular viewing tubes (inside diameter ⅞ in.) inserted in a large screen. Viewing distance was 36.5 in. and the *S*'s visual angle was approximately 17 deg. The stimulus surface was mounted on a board pivoted on its central horizontal axis. A chain-and-sprocket drive connected the board to a small wheel at *S*'s right hand. *E* changed the slant of the surface by moving the board itself. A protractor at one of the axis supports could be read to .5 deg. Two 20-w. fluorescent tubes provided nonuniform but constant illumination irrespective of changes in slant. These were mounted on the slanting board, not visible to *S*. A manually operated shutter occluded the surface between trials and during *E*'s resettings of the surface. Gross head movements were eliminated by an adjustable chinrest.

Subjects. The *Ss* were 13 students from an introductory course in psychology and 3 graduate students, 10 men and 6 women in all. Three other *Ss* were rejected (see Procedure).

Procedure. A preliminary study had shown that practice trials, without knowledge of results, significantly reduced the posttest adjustment time. Each *S* therefore made 20 practice adjustments to the vertical from six starting positions, which were randomly assigned with the restriction that −2, −1, +1, and +2 deg. each occurred four times, and −5 and +5 deg. each occurred twice.[10] During the test itself, each *S* was stimulated and tested twice, with a 2-min. inducing period and a 5-min. interval from end of stimulation on Trial 1 to the onset of the pretest on Trial 2. Half of the *S*s were stimulated with +30 deg. slant on Trial 1 and −30 deg. on Trial 2. The other eight *S*s were given the reverse order. On each trial there was a single pretest and a single posttest.

---

[9] With $\alpha = \beta = .05$ as Type I and Type II error rates respectively, repeated measures on $N = 6$ (subjects) allows detection of a true difference between two means of 1.85σ (W. C. Guenther, *Concepts of Statistical Inference*, 1965, 154). The smallest variance obtained by Bergman and Gibson gives an estimate of σ of 1.41 (W. J. Dixon and F. J. Massey, *Introduction to Statistical Analysis*, 1957, 76). Hence the estimated minimum detectable true difference between a pair of means is 2.61 deg.

Starting positions were −2, −1, +1, and +2 deg. Pretest and posttest starting positions were the same on any trial, but each *S* had a 'positive' starting position on one trial and a 'negative' starting position on the other. Each of the eight possible starting-position combinations was allocated to two of the *S*s. Of the two *S*s with the same starting-position combination, one was stimulated with +30 deg. first, the other with −30 deg. first.

Each *S* was instructed to fixate the central white line continuously during adjustment and stimulation, and to adjust the screen until it was physically vertical. He was told that the surface would never be vertical when first presented. Bracketing was partially controlled by instruction; *S* was instructed to 'overshoot' at least once and not more than once in each direction.[11]

One disadvantage of using a regularly contoured surface is the occurrence of the "wallpaper" phenomenon, or *Tapetenbilden*.[12] Diplopia occurs but is not reported in the absence of a fixation point, since the amount of under- or overconvergence is such that immediately adjacent contours exactly overlap. The *S*s were instructed to report immediately whenever *two* fixation lines were visible. Three *S*s who reported the effect continuously during practice and stimulation were rejected. In other cases, the effect was of brief duration and could often be removed by the introduction of a moving stimulus (e.g., a pencil) into *S*'s field of vision.

Measure of aftereffect. Aftereffect was determined by the correction method.[13] When the inducing stimulus is slanted equally in opposite directions on a consecutive pair of trials (balanced trials), incomplete dissipation of aftereffect induced on Trial 1 may affect the pretest of Trial 2. Such a carry-over effect would spuriously inflate the second aftereffect measure. Let the two directions of inducing slant be labeled + and −. If the pretests and posttests on these trials are respectively labeled A and B (for the + trial) and C and D (for the − trial), then $[(B - A) + (C - D)]/2 = [(B - D) + (C - A)]/2$ is a common measure of the mean aftereffect.[14] If there is no systematic carryover effect, then C and A may be regarded as estimates of S's 'true' judged vertical. Hence, the expected value of $(C - A)/2$ is zero and $[(B - D) + (C - A)]/2 \equiv (B - D)/2$. The correction method defines the aftereffect as $(B - D)/2$. In the absence of a carryover effect, the correction method and the more common (noncorrection) method give the same result. If the significant carryover effect occurs, then $[(B - D) + (C - A)]/2 > (B - D)/2$.

---

[10] When the top of the surface was nearer *S* than the bottom, the slant was called +; when the top was farther from *S* the slant was called −.

[11] A similar instruction was used by Bergman and Gibson (n. 3), 368.

[12] J. Le Comte, On some phenomena of binocular vision: Some peculiarities of the phantom images formed by binocular combination of regular figures, *Amer. J. Sci.*, 3d ser., 34, 1887, 97–107; H. J. Engel and M. H. Fischer, Optokinetic space perception: A new principle, *Pflug. Arch. f.d.g. Physiol.*, 233, 1950, 127–146; J. Beloff, The stripe paradox, *Brit. J. Psychol.*, 52, 1961, 323–331; I. Lie, Convergence as a cue to perceived distance, *Scand. J. Psychol.*, 6, 1965, 109–116.

[13] P. M. Wenderoth, A definition of after-effect to be used in the case of paired, balanced trials, paper read at the Sixth Canberra Symposium on Perception, Canberra, 1964; Bergman and Gibson (n. 3), 368.

[14] E.g., G. Singer and R. H. Day, Temporal determinants of a kinesthetic after-effect, *J. exp. Psychol.*, 69, 1965, 343–348.

Results

The mean aftereffect was 3.58 deg., with variance of 2.84. The effect was significant ($t = 8.23$; $t_{.99} = 2.947$). Mean time between the end of stimulation and the end of the posttest was 10.7 sec., with variance of 4.1. This mean time included an average 3.8 sec. for the experimenter to reset the stimulus.

A pilot study had shown that the starting position had a significant effect on pretests when starting-position slants were as large as ± 5 deg. This suggested that the slant should be reduced so that regression effects would be small relative to the size of the aftereffect. However, the effect of starting-position slant in the experimental pretests was not significant ($F = 1.84$; $F_{.95} = 2.82$).

Discussion

The results confirmed the prediction that a significant aftereffect could be induced by a 2-min. fixation of a high-contrast, regularly textured stimulus surface. The obtained mean was larger than those reported by Bergman and Gibson, though the variance was also relatively large.

Despite the 5-min. intertrial interval, there appeared to be a significant carryover from Trial 1 to Trial 2. The mean first measure was 5.25 deg. and the mean second measure was 3.47 deg. $\Sigma(C - A)/2n$ was .83 deg., with variance of 1.34. This was significant ($t = 2.73$; $t_{.95} = 2.13$).

EXPERIMENT II

The aim of this experiment was to examine the dissipation of the slant aftereffect by measuring the effect at various intervals after the end of stimulation. It has been claimed, in the case of other spatial aftereffects, that the dissipation function is exponential.[15] An exponential decay function was therefore fitted to the means of this study and a test of goodness of fit was carried out. It was also predicted that the aftereffect would be zero 5 min. after stimulation offset.

---

[15] M. Sagara and T. Oyama, Experimental studies of figural after-effects in Japan, *Psychol. Bull.*, 54, 1957, 327–338; Singer and Day (n. 14).

Method

Apparatus. The apparatus was identical to that of Experiment I.

Subjects. The *S*s were students from an introductory course in psychology, seven men and seven women. No *S*s wore spectacles.

Procedure. Each *S* was given a pair of balanced trials under each of seven treatment conditions. One of each pair of trials was with an inducing-stimulus slant of +30 deg., the other with −30 deg. Posttests were commenced 0, 30, 60, 90, 120, 150, and 300 sec. after the end of stimulation.[16] Each *S* had six trials (three treatment conditions) on Day 1 and eight trials on Day 2. Each session took approximately one hour and was preceded by 10 practice settings. The order of treatments over *S*s was partially counterbalanced so that each delay was given twice in each order. The interval between the end of stimulation and the start of the next pretest was always 5 min. A posttest always occurred in this 5-min. interval.

The starting positions were the same as in Experiment I except that counterbalancing of the eight starting-position combinations over subjects and treatments was not possible. Six starting-position combinations occurred twice and the remaining two once each.

Results

Means and variances of aftereffects measured following the seven posttest delay intervals are shown in Table I. An exponential function was fitted to six of the groups—the 300-sec. condition excluded—and the predicted means are shown in Table I.[17] The exponential and residual contrasts among the obtained means were tested for significance by a method of unplanned tests.[18] The results of this analysis are shown in Table II. The exponential function accounted for the significant proportion of the variance among the group means. The residual variance was not significant. When the fitted exponential function was extrapolated to the 300-sec. post-

---

[16] These intervals do not include the mean 3.8 sec. for *E* to reset the stimulus following stimulation offset.

[17] J. P. Sutcliffe, An approach to the analysis of monotonic trends in data, unpublished paper, University of Sydney, 1964.

[18] R. S. Rodger, Type I errors and their decision basis, *Brit. J. math. statist. Psychol.* (in press). Values of $F_{(1-\alpha)}^{\nu_1}$ from R. S. Rodger, *Intermediate Statistics*, 1965. The method of analysis and statistical model are given by B. J. Winer, *Statistical Principles in Experimental Design*, 1962, 122–132.

TABLE I

Mean Aftereffects, Variances, and Predicted Means at Seven Posttest Delay Intervals: Experiment II

Measure	Posttest delay (sec.)						
	0	30	60	90	120	150	300
M	4.14	2.75	2.16	1.20	1.46	1.14	.93
S^2	1.93	1.11	.81	1.75	1.41	1.32	1.32
Y	4.15	2.76	1.97	1.53	1.29	1.15	.98

Note: All aftereffects, variances, and predicted means in degrees.

stimulus condition, the predicted value of .98 deg. was close to the obtained mean of .93 deg., which was significantly greater than zero ($t = 2.91$; $t_{.95} = 1.771$).

Discussion

This experiment indicated that the dissipation function of the aftereffect could be adequately fitted by an exponential curve, a finding which is consistent with results reported in other aftereffect studies. Although the aftereffect 5 min. after cessation of stimulation was significant, it was considered impractical to increase the intertrial interval. Hence, the subjects were tested once only in further studies.

EXPERIMENT III

Experiments III and IV were concerned with the ability of subjects to judge the slant of textured surfaces. Reports by the subjects and observations by the experimenter had suggested that

TABLE II

Summary of Analysis of Variance, Goodness of Fit of Exponential Function: Experiment II

Source	df	SS	MS	F_{obt}	$F_{.95}{}^{5}$	Decision
Between Ss	13	49.60				
Within Ss	70	163.20				
Exponential	5	91.64	18.33	17.29	1.44	Reject H_0
Residual	5	2.52	.50	.47	1.44	Accept H_0
Error	65	69.04	1.06			
Total	83	212.80				

the lattice pattern used in Experiments I and II was not altogether satisfactory for aftereffect measurement. Movement of the surface following prolonged fixation produced illusory movements similar to those reported by Honisett and Oldfield.[19] These made vertical adjustments difficult. In addition, the number of intersecting contours above and below the fixation point could have provided an unwanted cue to slant.

The aim of Experiment III was to measure the judged slant of two surfaces. One surface had vertical contours only, the other had only horizontal contours. It was predicted that binocular judgments of the slanted vertical contours would be relatively accurate, since these contours, when slanted, would provide the necessary condition for stereopsis, that is, discriminable horizontal disparities. On the other hand, since the edges of the targets were occluded and the fixation point was at the axis of rotation, the horizontal contours would not provide discriminable horizontal disparities. In that case, the disparate points would lie on continuous undifferentiated contours. If these predictions were confirmed, the horizontal contours could be omitted from the stimulus used to induce slant aftereffect without detracting from the binocular depth cues available.

Method

Apparatus. The apparatus was housed in a wooden box. The interior of the box was divided by a diagonal septum into which there was set an 8-in. square semisilvered mirror. The variable surface (test stimulus) was viewed through the mirror, the standard (inducing stimulus) was reflected. Both stimuli were mounted on large aluminium sheets and held in place by bulldog clips.

The viewing distance for both figures was approximately 28.5 in. The *S* observed the stimuli through adjustable-convergence tubes, 1.6 in. in diameter. All *S*s reported complete or almost complete overlap of left- and right-eye fields when the proximal ends of the tubes were set 2.6 in. (66 mm.) apart. The tubes occluded the edges of the stimulus surfaces and permitted a view of 10 deg. 50 min. in the center of each field.

The slant of the variable test figure was controlled either by *E* or by *S*, using small brass wheels which were connected to a 12-in. protractor at the rear of the box. Only *E* could vary the slant of the inducing figure.

The test figure was illuminated by two 25-w. bulbs, mounted in housings on each side of the backing plate. The inducing figure was similarly illuminated by 60-w. bulbs. Brightness of both fields was 40–50 ftL. at the center but increased rapidly to 100 ftL. at the periphery. The test and inducing figures could be alternately presented using decade timers or by a manual control switch.

---

[19] J. Honisett and R. C. Oldfield, Movement and distortion in visual patterns during prolonged fixation, *Scand. J. Psychol.*, 2, 1961, 49–55.

Three stimulus surfaces were used in this experiment. The surface of the variable test figure was plain white 1/16-in. pasteboard with four 1-in. squares of the same board pasted symmetrically around the central fixation point. The standard inducing figures, either horizontally or vertically contoured, were photographs mounted on pasteboard. In each case, 15 black lines, 3 mm. wide, were separated by white spaces, 12.5 mm. wide, with a central fixation spot.

Subjects. The *S*s were 30 students from an introductory psychology course.

Procedure. The *S*s were randomly assigned to one of two groups. Group V judged the vertical contours and Group H the horizontal contours. Each group was further divided into five groups of three *S*s. Each of these groups judged the standard at one of five slants: 0, −10, −20, −30, or −40 deg.

The staircase method was used.[20] *S* always judged whether the variable was 'more' or 'less' slanted than the standard, but the order of presentation of the standard and variable was alternated.

Each *S* completed two single staircases, each staircase consisting of five blocks of 10 trials. The slant of the variable, which was initially either +8 or −40 deg., changed in steps of 8 deg. until 10 consecutive responses alternated between two values. The next trial began at the center of these two values, with the step size reduced to 2 deg. The first trial of each series was designed as the trial preceding the first reversal in response.

One trial in each block of 10 was a catch trial. This was three or four steps above or below the preceding response. To reduce drift due to adaptation, a 1-min. rest was given between blocks of 10 trials and there was a 5-min. rest between the two series. To reduce the variance of judgments, *S*s were told that the standard slant would be constant over a single staircase. *S* was instructed to fixate the center of each surface at all times.

Results

The approximate formula of Dixon and Massey was used to estimate the means of the two staircase series (n. 9; pp. 318–327). These were similar in all cases and were combined. Table III shows the means and variances for each of the experimental groups. There were large variance differences, but Cochran's test suggested that the assumption of homogeneity was tenable.[21] Inspection of the data suggested tests of the linear and residual regression components of Groups V and H, and a test of the difference between the general means of the two groups. Rodger's method of unplanned tests was used (n. 18). In Group H, there was no significant linear

[20] D. J. Finney, *Probit Analysis,* 1962.

[21] See Guenther (n. 9), 219.

TABLE III

MEANS AND VARIANCES OF MEAN PSEs AT EACH SLANT ANGLE FOR TWO STANDARD SURFACES: EXPERIMENT III

Standard contours	Measure	Slant of standard (deg.) 0	−10	−20	−30	−40
Vertical	M	− .66	−6.45	−10.84	−17.96	−21.76
	S^2	.24	2.23	32.25	13.51	16.76
Horizontal	M	−3.19	−3.34	− 3.77	− 2.75	− 2.14
	S^2	10.09	.06	.14	8.43	6.54

Note: $N = 3$ in each cell.

or residual trend ($F = .02$ and $.02$ respectively; $F_{.95}{}^{9} = 1.16$). In Group V, the linear trend was significant ($F = 7.27$) and was fitted by the equation $Y = .537x + .79$. The residual trend was not significant ($F = .05$). The general means of Groups V and H differed significantly ($F = 4.55$).

Discussion

The results confirmed the prediction that binocular judgments of the vertical contours would be relatively accurate, while those of the horizontal contours would not. These results agree with those of Ebenholtz and Walchli, who postulated an inhibitory mechanism to account for loss of stereopsis.[22] It is sufficient to recognize, however, that the disparities would not be discriminable along homogeneous contours.

EXPERIMENT IV

Bergman and Gibson found no difference between monocularly and binocularly induced slant aftereffects (n. 3; p. 371). Bergman argued that a difference between monocular and binocular posttest settings could either be a function of a difference in the size of aftereffects induced by monocular and binocular stimulation or solely a function of differences between test adjustments under the two viewing conditions. However, he claimed that no difference between test judgments would indicate no difference between induced aftereffects.[23] This argument is a non sequitur, since it ignores the possibility of interaction effects. The addition of two further conditions is necessary: binocular stimulation with monocular test, and monocular stimulation with binocular test.

---

[22] S. M. Ebenholtz and R. M. Walchli, Stereoscopic thresholds as a function of head- and object-orientation, *Vis. Res.*, 5, 1965, 455–462.

[23] R. L. Bergman, After-effects in the perception of a slanted surface, unpublished M.A. thesis, Cornell University, 1954, 21–22.

The aim of this experiment was to compare the means and variances of vertical adjustments of the vertically contoured surface under monocular and binocular viewing conditions. If the variances were comparable, differences between aftereffects induced under the two viewing conditions could be tested.

Method

Apparatus. The test stimulus was the vertically contoured surface of Experiment III; no inducing stimulus was used. Either viewing tube could be occluded by covering the end inside the box.

Subjects. The *Ss* were 10 students from an introductory course in psychology; none wore corrective lenses.

Procedure. Each *S* was given a standard test of ocular dominance (visual alignment), which provided an objective criterion for selection of the test eye. The five *Ss* in Group BM made 10 binocular, followed, after a 1-min. rest, by 10 monocular, adjustments to the vertical. Group MB made the monocular adjustments first. *Ss* fixated the center of the surface at all times. Starting positions alternated between +2 and −2 deg., and were the same for all four conditions.

Results

Means and variances of vertical adjustments are shown in Table IV. The mean variance of the monocular judgments was significantly larger than that of the binocular judgments, both for first judgments ($F = 34.9$; $F_{.99} = 15.98$) and second judgments ($F = 32.1$).

Discussion

The results of Experiment IV suggested that the vertically contoured surface provided sufficient cues for relatively accurate binocular settings to the vertical, but not for accurate monocular settings. Since monocular mean settings ranged from −11.05 to +23.60 deg., detection of an aftereffect of the order of 4 deg. under these conditions would have required a prohibitive sample size. Thus binocular viewing only was used in Experiment V.

EXPERIMENT V

The object of this experiment was to measure the relationship between inducing-stimulus slant and the size of the induced after-

TABLE IV

Means and Variances of 10 Vertical Settings under Monocular and Binocular Viewing Conditions: Experiment IV

	Viewing condition			
	Binocular		Monocular	
Subject	*M*	S^2	*M*	S^2
Group BM	+1.65	1.30	+2.43	130.90
1	− .15	.76	+ .75	25.32
2	+1.60	12.19	−11.05	27.63
3	+1.55	2.53	+23.60	37.89
4	+1.80	2.46	−1.95	12.93
5	+3.43	7.58	+ .80	86.11
Group MB	+2.88	4.07	+4.32	45.59
6	+4.75	12.02	−2.60	27.64
7	+3.85	7.71	−3.65	112.56
8	+1.50	5.60	+6.30	83.56
9	+1.10	3.94	+6.95	359.98
10	+3.20	4.56	+14.60	81.14

effect. Each *S* was tested once, and only negative inducing slants were used. An attempt was made to obtain judgments of inducing-stimulus slant.

Method

Apparatus. The apparatus was the same as that of Experiment IV, except that an inducing stimulus identical to the test stimulus was used. In addition, a fixation light was mounted behind a pinhole drilled through the center of the test stimulus. Each *S* judged inducing-figure slant by adjusting the tilt of a black rod pivoted at the center of a white disc.

Subjects. The *Ss* were 70 students from an introductory course in psychology, 30 men and 40 women. None wore corrective lenses. A total of 20 other *Ss* was rejected: 13 *Ss* had a pretest variance which exceeded 16.0, 4 *Ss* reported continuous double images of the fixation point despite use of the fixation light (see Procedure), and 3 *Ss* could not discriminate large differences in slant of the test stimulus.

Procedure. There were 10 *Ss* in each group. Groups were defined by inducing-stimulus slants: 0, −10, −20, −30, −40, −50, and −60 deg. Each *S* was instructed to fixate the pinhole of light on the test stimulus before that figure was illuminated. This procedure helped to reduce the occurrence of convergence 'locking.' Ten pretest settings were made from starting positions alternately +2 and −2 deg. The *S* then inspected the inducing stimulus for 2 sec. before he adjusted the metal rod to match the inducing-stimulus presentations. This was repeated twice, with a 1-min. rest between inducing-stimulus presentations.

Ninety sec. after the last slant judgment *S* made one vertical setting of the test stimulus, and then immediately fixated the inducing stimulus for 90 sec. After 90 sec. had elapsed, *S* was instructed to set the test stimulus, which was set at 0 deg., to physical vertical.

Results

The means and variances of aftereffects measured at each inducing-stimulus slant are shown in Table V. Cochran's test indicated that the assumption of homogeneity of variance was untenable. Correlation between sample means and variances appeared not to be the major cause of inhomogeneity ($r = +0.40$, $t = 1.00$, $p > .05$). There was, however, a significant correlation between aftereffect variance and inducing-stimulus slant ($r = +0.94$, $t = 6.064$, $p < .01$). There was also a consistent tendency for the variance of slant judgments to increase with physical slant (from 59.28 at 10 deg. to 128.18 at 60 deg.). Illusory movement effects occurred despite the absence of horizontal contours. Some subjects reported these effects during the posttest, others when subsequently questioned.

TABLE V

Means and Variances of Aftereffects
Measured at Seven I-figure Slants: Experiment V

	Inducing-stimulus slant (deg.)						
Measure	0	−10	−20	−30	−40	−50	−60
M	−1.72	1.48	2.65	7.11	2.89	4.22	3.39
S^2	9.40	9.50	11.09	19.14	31.54	28.36	44.32

Discussion

The results of this experiment were inconclusive. The illusory movement effects could be related to retinal tilt of afterimages of the inducing-stimulus contours relative to the test-stimulus contours. Such relative tilt would increase with inducing-stimulus slant. Thus, contour-interaction effects could explain the correlation of aftereffect variance with inducing-stimulus slant. The increase in variance of judged slant with increases in inducing-stimulus slant provides an alternative possible explanation, which is discussed below.

CONCLUSIONS

The results of these preliminary experiments suggested that the present approach to the study of slant aftereffect was unsatisfactory. A textured *surface* was chosen partly to enable a comparison between monocularly and binocularly induced effects. The occurrence of illusory movements and the large variability of monocular adjustments to the vertical suggested that the surface chosen was not adequate for such a comparison. Monocular slant judgments might have increased in accuracy if the 'interaction' of texture elements had been increased by decreasing texture coarseness.[24] However, the experiments indicated that this would not be a suitable alternative, because illusory effects and *Tapetenbilden* interfered with posttest judgments. Bergman and Gibson's alternative of using a low-contrast surface was rejected because a very long inducing period results in very small effects, thus making differences between treatments difficult to detect when sample size is small. There were other reasons for preferring to work with a textured surface, including possibilities suggested by Köhler and Emery (n. 2; pp. 173–174). For example, it would have been interesting to compare aftereffects induced by a slanted surface and by a nonslanted pattern of converging vertical contours.

The decision not to measure monocularly induced slant aftereffect suggests that a single vertical contour could be used as both inducing stimulus and test stimulus. The monocular cue of linear perspective is not necessary for binocular depth perception. In further experiments, therefore, a single vertical contour is being used.[25] Such a stimulus does not induce either illusory movement or the *Tapetenbilden* phenomenon, both of which occurred in Experiment V.

Recent evidence strongly suggests that both Gibson's 'normalization' (adaptation) theory and Köhler's contour-shift, or repulsion, account are by themselves inadequate as explanations of all aftereffect phenomena.[26] Nevertheless it is interesting to note that an implication of Gibson's theory has relevance for the results of Experiment V. The adaptation theory is formulated in phenomenological terms. Thus slant aftereffect would be predicted to be a function of judged rather than physical slant.[27] Since there was some tendency for the variance of judged slant to increase with actual slant, Gibson's theory would predict that aftereffect variance would also increase with judged slant. An adequate test of this prediction would require more precise methods of measuring judged slant than were employed in Experiment V.

---

[24] H. E. Gruber and W. C. Clark, Perception of slanted surfaces, *Percept. mot. Skills*, 6, 1956, 97–106; E. S. Eriksson, Monocular slant perception and the texture gradient concept, *Scand. J. Psychol.*, 5, 1964, 123–128.

[25] P. M. Wenderoth, R. S. Rodger, and I. S. Curthoys, Confounding of psychophysical errors and sensory effects in adjustment measures of spatial aftereffects, *Percept. & Psychophys.*, 4, 1968, 133–138.

[26] Some of the more recent studies are R. B. Morant and J. R. Harris, Two different after-effects of exposure to visual tilts, this JOURNAL, 78, 1965, 218–226; J. Wilson, Adaptation and repulsion in the figural aftereffect, *Quart. J. exp. Psychol.*, 17, 1965, 1–13; W. B. Templeton, I. P. Howard, and G. Easting, Satiation and the tilt after-effect, this JOURNAL, 78, 1965, 656–659; L. Ganz and R. H. Day, An analysis of the satiation-fatigue mechanism of figural aftereffects, this JOURNAL, 78, 1965, 345–361.

An alternative hypothesis of the determinants of slant aftereffect could be tested more easily with a single slanted contour, used as both inducing and test figure, and similar to the stimulus used by Ogle to study cyclofusional eye movements.[28] Variation in the slant of a vertical contour in an observer's median plane is projected retinally as variation in tilt relative to the vertical retinal meridian. There is no essential difference, therefore, between prolonged monocular fixation of a very thin contour tilted through an angle $\Theta/2$ and a contour slanted through an angle $\Theta' = \tan^{-1}[(b/2a)\tan\Theta]$, assuming that cyclotorsional eye movements do not occur.[29] Thus binocular slant aftereffect may be a complex case of tilt aftereffect in which directionally opposite tilt effects are induced at the two eyes.

---

[27] Gibson (the first reference in n. 5).
[28] K. N. Ogle, *Researches in Binocular Vision,* 1950, 107.
[29] Ogle (n. 28), 103.

TRACE LOCALIZATION AND THE RECOGNITION OF VISUAL FORM

CHARLES C. TORREY, *Macalester College*

Abstract. In the first of three experiments, previous research by Wallach and Austin suggesting localized storage of visual traces was replicated, using new stimulus materials. In the second, an alternative hypothesis accounting for trace selection by phenomenal, rather than anatomical, locus was examined and rejected. In the third, the possibility that eye-movement tendencies may provide mediating stimuli that govern trace selection was explored, with negative results. The hypothesis of differential trace localization originally put forth by Wallach and Austin remains the most economical explanation of their results, but neither their research nor the present experiments indicate whether the localized trace is best described as a template or as a set of features.

In 1954, Wallach and Austin reported an experiment on visual recognition the results of which, they believed, indicated that memory traces of visual stimuli are localized in a cortical region specific to the region of original excitation.[1] The experiment hinged upon the use of a silhouette figure which when viewed in one orientation represented a man wearing a chef's hat, but when rotated 90 deg. to one side represented a dog. The chef, the dog, and seven other silhouettes were shown one at a time tachistoscopically, and the observer was required to identify them. Each figure appeared in one of four quadrants of the projection screen, with the observer fixating the center point during the exposure. The two versions of the critical figure were shown second and sixth in the series, one to the upper left and one to the upper right of the fixation point. As a tenth and final figure, the critical figure was again presented in either the upper left or upper right quadrant, but oriented diagonally so as to constitute an 'ambiguous' figure.

Of the 48 observers who completed the experiment, 36 responded

---

Received for publication by Professor E. B. Newman. The author thanks Steven J. Taff for his assistance in conducting the experiments and Richard V. Schultz for help in preparing slides for Experiments I and II.

[1] Hans Wallach and Pauline Austin, Recognition and the localization of visual traces, this JOURNAL, 67, 1954, 338–340.

to the ambiguous figure as they had earlier responded to the critical figure presented in the same quadrant. This frequency of *near-plant* responses differed significantly from chance ($p < .002$), and the authors concluded that trace selection had been influenced by anatomical proximity between the excitation and the trace of the earlier excitation.

It has long been accepted as fact that the recognition of a figure is unimpaired when that figure is presented at a novel locus on the retina.[2] Because Wallach and Austin's data suggest that anatomical locus can, under some circumstances, influence the trace-selection stage of visual recognition, it seemed important to replicate their experiment and to explore some possible alternative explanations of their results. The present series of experiments was undertaken for these reasons.

EXPERIMENT I

Method

The stimulus forms were black silhouettes of familiar animals, including a squirrel, a mouse, an owl, a whale, a turtle, a snail, and a pig. The critical figure represented a duck (D figure) when shown in one orientation and a rabbit (R figure) when rotated 90 deg. The tenth figure of the series (Figure 1) was the duck-rabbit silhouette in an intermediate orientation (A figure). The 35-mm. slides prepared for each figure were appropriately placed in the field, and the figures were projected on a screen at a distance of 54 in. from

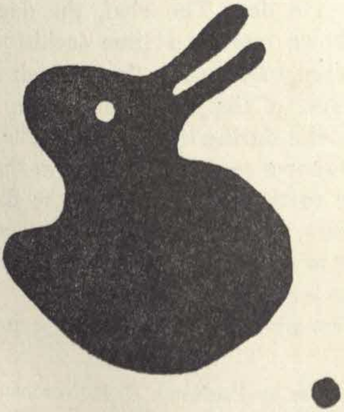


FIG. 1. Ambiguous (A) figure used in the experiments

---

[2] Wolfgang Köhler, *Dynamics in Psychology*, 1940, 127–129.

O. The long dimension of the projected figures ranged from 3.5 to 4.5 in., and the figures were centered at points to the upper left, upper right, lower left, or lower right of the thumbtack that served as fixation point, at a distance of from 2.5 to 3 in. from the thumbtack. In all cases, the nearest portion of the figure contour was separated from the fixation point by at least 1 deg. visual angle, and thus lay outside the foveal region. Exposure of each figure was controlled at .4 sec. by means of a camera shutter mounted in front of the projection lens. The experiment was conducted in a black-painted room that was dimly illuminated by a fluorescent ceiling fixture during the experiment.

Each *O* was seated facing the screen and was given the following instructions: "This is an experiment on how people recognize things. I will flash some pictures on the screen very quickly. Looking directly at the thumbtack, will you name the object which you see? Remember, always look directly at the thumbtack during the exposure. I will give you a ready signal before each picture. Any questions?"

The D and R figures were shown second and sixth in the series, one to the upper left and one to the upper right of the fixation point. The A figure was presented last, in either the upper left or upper right quadrant. Order of presentation for D and R figures was counterbalanced, as was position relative to the fixation point for D, R, and A figures. The other seven figures were presented, in various quadrants, first, third through fifth, and seventh through ninth in the series, in a constant order and position for all *O*s.

A total of 65 undergraduate students from the introductory psychology course at Macalester College served as *O*s. The eight presentation sequences generated by the counterbalanced design were used in rotating order from one *O* to the next, and if in the course of the experiment *O* failed to identify the D or R figure, made the same response to both, or made a completely new response to the A figure, his data were excluded from further consideration and the next *O* was run under the same sequence. On these grounds 17 *O*s were dropped from the experiment, and of the remaining 48, 6 were run under each stimulus sequence.

Results

Of the 48 observers, 36 identified the A figure in the same way they had earlier responded to the critical figure presented in the same position (near-plant), and 12 gave the other (far-plant) response. This distribution of response differs significantly from chance ($p < .002$). It is interesting that the proportion is identical to that obtained by Wallach and Austin, using a different critical figure.

Wallach and Austin reported that of the 12 far-plant recognitions in their experiment, in 9 cases the effective plant had been made in the left cerebral hemisphere. They speculated that hemispheric dominance might play a role in recognition, over and above the effect of trace locus. The present data did not support this hypothesis, since in 9 of the 12 cases of far-plant recognition, the effective plant was in the right visual cortex.

Furthermore, there was little indication in the present data of other factors influencing response to the A figure. Of the 48 responses, 28 were 'duck' and 20 'rabbit,' indicating that the A orientation was fairly well balanced between the D and R orientations. Also, 26 observers gave the same response to the A figure as to the first critical figure (Trial 2), and 22 the same response as to the second critical figure (Trial 6). There was thus no evidence of a primacy or recency effect. In sum, the results constitute a clear replication of the Wallach and Austin experiment, and demonstrate that the near-plant effect is not limited to the particular critical figure used there.

EXPERIMENT II

In Experiment I, as in the Wallach and Austin experiment, two distinct kinds of localization were confounded. Not only was the excitation process localized anatomically, with reference to the retina and to the visual projection area, but the silhouette was localized with reference to the projection screen. It is thus conceivable that the operative factor in determining trace selection in these experiments was not anatomical location but 'phenomenal' location. Some indirect support for the plausibility of this hypothesis can be found in a series of experiments by Rock, in which the recognition of simple figures was shown to be more influenced by their phenomenal orientation than by their orientation with respect to retinal coordinates.[3] The second experiment of the present series adapted to the study of localization one of Rock's methods for separating retinal and phenomenal factors.

Method

Design. The experiment pitted the two types of localization, anatomical and phenomenal, against each other so as to assess their relative influence on recognition of the A figure. By shifting the fixation point appropriately, it was possible to present the A figure at a screen position earlier occupied by the D figure, but at the same time to that part of the retina, and hence to that hemisphere of the cortex, where the R figure had been presented earlier. The relative importance of anatomical and phenomenal reference systems could thus be determined by comparing the frequency of 'rabbit' and 'duck' responses to the A figure. To control for possible effects due merely to the frequent shifts of fixation, a second group was run with shifting fixation, but with phenomenal and anatomical locations deliberately confounded as in the earlier experiments.

---

[3] Irvin Rock, The orientation of forms on the retina and in the environment, this JOURNAL, 69, 1956, 513–528; Irvin Rock and Walter Heimer, The effect of retinal and phenomenal orientation on the perception of form, this JOURNAL, 70, 1957, 493–511.

Procedure. Two fixation points, a red and a black thumbtack, were permanently attached to the screen, separated horizontally by a distance of 5 in. The silhouettes were those used in Experiment I, save that a giraffe was substituted for the snail, which many *O*s had failed to recognize.

Each *O* was seated at a distance of 54 in. from the screen and the following instructions were read: "This is an experiment on how people recognize things. I will flash some pictures on the screen very quickly. Before I flash each picture, I will tell you which of the two thumbtacks to look at, and then give you a ready signal. Looking directly at the thumbtack, will you please name the object which you see? Please try to keep your body position fairly constant and remember to look directly at the thumbtack during the exposure. Any questions?"

Each figure of the series was presented in one of the four quadrants immediately adjacent to the fixation point for that trial. This necessitated making duplicate slides for many of the noncritical figures, and triplicate slides for the D and R figures with differences only in location of the figure within the field. The A figure was always presented above and between the two fixation points (left visual field for *O*s fixating the right thumbtack). For each *O* the critical figures were presented on the second and sixth trials, to the upper left and upper right of the same fixation point. The only procedural difference between the experimental and control groups was that on the tenth trial the control *O*s fixated the same point as on the second and sixth trials (anatomical and phenomenal locations confounded), while the experimental *O*s fixated the other point (anatomical and phenomenal locations separated). Since all *O*s were required to shift fixation points frequently (six times for experimental *O*s, seven times for control *O*s), the possible influences of such fixation shifts were approximately equal for the two groups.

The dimensions of the projected figures, their positions in relation to the fixation point for that trial, and all other conditions of room illumination, exposure duration, and so on were the same as in Experiment I. For both groups, order and position of D and R figures, and fixation point on the tenth trial, were counterbalanced, and the 16 sequences (8 control and 8 experimental) were run in rotation as in Experiment I, until each experimental sequence had six usable cases and each control sequence four. In all, 85 naive *O*s, obtained from the same source as those of Experiment I, took part, and of these 5 were eliminated from the experiment because of recognition failures of the sorts described above.

Results

Of the 48 observers in the experimental group, 37 responded to the A figure in the same way they had earlier responded to the critical figure presented at the same anatomical position, while 11 gave the same response as that elicited by the critical figure presented at the same screen position. Since this response distribution differs significantly ($p < .001$) from what would be expected if the two factors were equally effective, it is clear that phenomenal location was the less potent factor in determining trace selection.

As in Experiment I, there was no evidence of a primacy or recency effect in either of the present groups, nor was there any indicated predominance of 'duck' or 'rabbit' responses to the A figure. In the experimental group, the 11 far-plant recognitions were examined for evidence of hemispheric dominance, again with negative findings: in 6 of these cases the effective plant was made in the left cortex; in the other 5, in the right cortex.

In the comparison of experimental and control groups, however, there was a suggestion of some minor influence from phenomenal location, in that for the latter group only 3 out of 32 observers gave far-plant responses to the A figure. While the difference (between 11/48 and 3/32) was not significant ($p > .05$, one-tailed test for difference between uncorrelated proportions), the direction was appropriate to the hypothesis that phenomenal location can slightly inflate or reduce other influences on trace selection, depending upon whether it is confounded with, or set in opposition to, those influences. Nonetheless, there remained in the experimental group a significant proportion of anatomically near-plant responses, which cannot be explained by appeal to phenomenal location.

EXPERIMENT III

Seeking alternatives to Wallach and Austin's explanation of their results, Neisser suggested two hypotheses.[4] The first of these is that location of the figure within the visual framework may be influential in the selection of traces. Experiment II seemed to show that such influence is at best minimal. As a second possibility, Neisser suggested that response selection may be mediated by "specific feelings of eye-strain," presumably produced by tendencies to move the eyes toward the side where the figure appears. Thus, eyestrain evoked by a figure to the left of fixation—say, a figure identified as a 'duck'—might, when again evoked, mediate the response 'duck' to the somewhat similar A figure.

It does not appear possible to rule out this hypothesis through direct experimental test, since any peripheral presentation of a figure (essential to the 'planting' of traces in differential cortical loci) inevitably evokes reflex tendencies to shift the eyes in that direction. Eyestrain as a potential mediating stimulus is thus invariably confounded with differential anatomical locus of the visual excitation. Even the use of a stabilized-image technique, though it would render futile an observer's attempts to fixate the figure, could not be expected to eliminate the reflexive elicitation of eye-movement tendencies.

---

[4] Ulric Neisser, *Cognitive Psychology*, 1967, 52–53.

If mediation does play an important role in trace selection under the conditions of the Wallach and Austin experiment, however, it should be possible to show the same selective effects with foveal (rather than peripheral) presentation, provided the observer has been given appropriate mediating stimulation. It does not seem that *suppression* of eye movements, as presumably occurs under instructions to maintain constant fixation, is an important component of the suggested mediating mechanism. If the selective effect depends on the specification of differential direction from the fovea by a mediating process involving the extrinsic muscles and possible sensory feedback therefrom, then selection should occur whether or not the eyes are actually shifted to the position of the figure just prior to its identification by the observer. Further, the use of verbal mediators (e.g., instructions to shift fixation to the 'upper left' just before the figure appears there) should perhaps enhance the effect. On the other hand, if differential storage location is responsible for trace selection, as Wallach and Austin suggested, and if mediation by eye movements plays no role, then when the observer views both of the D and R figures as well as the A figure foveally, his response to the latter should revert to chance.

Accordingly, in Experiment III two groups were run as follows: both fixated the center at the beginning of each trial, and both were told the screen position in which the figures would be presented. The experimental group was instructed to shift fixation to the position indicated, and thus viewed the figure foveally when it appeared. The control group was instructed to maintain central fixation throughout exposure of the figure.

If the hypothesis of mediation by eyestrain is correct, both groups should show a significant predominance of near-plant responses to the A figure, since both the verbal specification of screen position and the ensuing sensory feedback from the extrinsic muscles provide potential mediating support for that response. If the theory of cortically localized traces is correct, the experimental group should show a chance relation between the positions of the D and R figures and subsequent response to the A figure, since all plants were actually foveal, while the control group should continue to give a predominance of near-plant responses to the A figure.

Method

The experiment was conducted using a two-field Dodge tachistoscope (Polymetric V-0959). The visible field between exposures was a buff-colored card on which was drawn a black grid consisting of an outline rectangle 3.5 in. wide × 3 in. high, divided into four quadrants by vertical and horizontal lines through its center. Silhouettes of the same animals as those used in Experiment II were drawn in black on buff-colored stimulus cards. The position of each figure on its card corresponded to the center on an appropriate quadrant of the preexposure field. The length of the silhouettes ranged from 1.0 to 1.25 in. and the nearest portion of the contour in all cases was at least .25 in. from the center of the grid (50 min. visual angle at the image distance of 17 in.).

All *O*s received these instructions: "This is an experiment on how people recognize things. In front of you is a grid of four rectangles. I am going to flash some pictures very rapidly in the positions indicated by the rectangles. I will say 'Ready; lower left,' and then I will flash a picture in the lower left quadrant. Or I will say 'Ready; lower right,' and flash a picture there." The experimental *O*s were also told: "When I say 'Ready,' look directly at the center of the grid. When I tell you the quadrant, shift your eyes to that quadrant so that when the picture appears you are looking at the proper part of the screen. When the picture has gone, tell me what it was. Remember to look at the center whenever I say 'Ready,' and then to look at the proper quadrant before the picture is shown. Any questions?" The control *O*s were told, for their supplemental instructions: "When I say 'Ready,' look directly at the center of the grid and continue looking that way during exposure of the picture. I am only telling you which quadrant so that you won't move your eyes that way when the picture appears. When the picture has gone, tell me what it was. Remember to keep your eyes fixed on the center of the screen during the presentation of each picture. Any questions?"

Apart from these changes in apparatus and instructions, the procedure was the same as that of Experiment I: order and position of presentation for the D and R figures, and position of the A figure, were counterbalanced within each of the two groups. Experimental and control conditions, and the sequences within each, were again run in rotating order. The exposure duration was .4 sec.

The 120 new *O*s were obtained from the same source as in the earlier experiments. There were 24 *O*s eliminated from the experiment because of recognition failures, and the remaining 96 were run in equal numbers under the experimental and control conditions.

Results

Within the experimental group, only 27 of the 48 observers gave the same response to the A figure as that they had earlier given to the D or R figure presented in the same screen position—and hence associated with the same eye movement. The number did not differ significantly from the 24 to be expected by chance. Of the 48 control observers, 35 gave near-plant responses, a frequency that did differ significantly from chance ($p < .003$).

The difference in frequency of near-plant responses between the two groups was significant at the .05 level, using a one-tailed test ($p < .044$). The one-tailed test was justified here in that a difference in the opposite direction between the groups would be difficult to interpret under either of the hypotheses being considered.

DISCUSSION

The results of Experiment III provide no support for the notion that the selection of traces in Wallach and Austin's recognition paradigm is mediated by "specific feelings of eyestrain." Admittedly the test of this hypothesis was indirect, but it appears valid, at least until the eyestrain hypothesis is further refined and until some evidence is provided that the suppression of eye movements is a necessary condition of their utility as mediating events.

The experiment might be criticized for its failure to provide closer control over the observer's actual eye movements than was possible through instruction. It may have been the case that some observers in the control group did shift fixation when the figure appeared, despite instructions; but even assuming that this occurred, it does not seem possible to so account for the overall pattern of results. If the selection of response to the A figure depends upon differential cortical location of the traces of the D and R figures, then any eye movements occurring when those figures were first presented should decrease the distance between their traces and accordingly reduce the predominance of near-plant responses to the A figure. If trace selection does not depend upon differential cortical locus, then the difference between the two groups in Experiment III must be attributed to a difference in mediating capacity between eye movements performed and eye movements suppressed, and the burden of proof is shifted to those who would suggest this distinction (as argued above). Considering the results of all three present experiments, as well as those of Wallach and Austin's experiment, it seems more plausible to assume that if anatomical position of stimulation were more carefully controlled, the proportion of near-plant responses might exceed the 75% that has remained so stable throughout all of these experiments.

While the results of all four experiments do suggest that traces of visual stimulation are, for a time at least, localized in a region of the cortex specific to the site of original excitation, no particular model of *what* is locally stored seems to be implied. Either a 'template' model or a 'features' model[5] of visual memory seems consistent with the notion of localized storage. Perhaps a feature interpretation is slightly favored, if only because template matching would seem to require some process whereby the A figure is 'rotated' by 45 deg. before its congruity with the near-plant can be established and because the direction of such rotation would already indicate the influence of the nearby trace. A 'part' feature such as 'two protuberances' is relatively invulnerable to changes under rotation and might well be established as 'ears of a rabbit' in one hemisphere and as 'bill of a duck' in the other.

---

[5] Neisser (n. 4), 47–52.

THE EFFECTS OF SIZE AND DISTANCE ON MAGNITUDE ESTIMATIONS OF APPARENT SIZE

ROBERT TEGHTSOONIAN and MARTHA TEGHTSOONIAN,
Smith College

Abstract. Size constancy in a natural viewing situation was studied by constructing scales of apparent area for irregular polygons at several viewing distances by the method of magnitude estimation (number matching). These scales had the form of a power relation, $\psi_A = kA^n$, where ψ_A is apparent area and A is physical area. For a stimulus of fixed area, ψ_A did not decrease with increasing distance, but increased (in Experiment I) or remained constant (in Experiment II). The relation between ψ_A and distance depended, first, on k, the scale factor, which, like ψ_A itself, increased with distance in Experiment I and remained constant in Experiment II, and second, on n, the exponent, which was invariant with distance in both experiments.

Of all the variables that might influence the apparent size of a stimulus, viewing distance has received the lion's share of attention. The paradoxical invariance of judged size despite the gross changes in retinal image produced by variation in viewing distance is the size-constancy effect, and it has generated a very sizable literature devoted to the search for reasons underlying the relative weakness of the distance variable in determining apparent size. One consequence of that interest, however, has been the neglect of other factors. Only recently has the effect of physical size, perhaps the most obvious determiner of apparent size, been scaled. Stevens and Guirao and M. Teghtsoonian found that only in the case of unidimensional stimuli is there support for the widespread assumption

---

Received for publication by Professor E. B. Newman. Parts of this paper were presented at the meetings of the Eastern Psychological Association, New York, 1966, as "Size constancy: Magnitude estimations as functions of stimulus area and distance." The data were collected at Smith College. The authors thank Anne Knecht Wilson, who made the stimuli, and Donna Greenough, who ran the subjects for Experiment I. This research was partially supported by Grant MH-11810-01 from the United States Public Health Service. The manuscript was prepared while the authors were visitors in the Behavioral Sciences Laboratory of the Bell Telephone Laboratories, Inc., Murray Hill, New Jersey. They thank Saul Sternberg for his discussion and his helpful comments on a draft of this paper.

that apparent and physical size are linearly related.[1] For two-dimensional stimuli, subjects judge in accord with the 'law of apparent area,' $\psi_A = kA^{.75}$. These results were obtained with the techniques of magnitude estimation and production, whose speed in use and ease of application made possible the consideration of a large range of stimulus values within a single experiment. This parametric approach stands in marked contrast to that of many experiments on size constancy which have used just one or two physical (or retinal) sizes. The typical size-constancy experiment has varied distance but not size, while the apparent-size experiment has varied size but not distance. Since in normal full-cue viewing, both size and distance vary over a large range of values, we wished to determine how these two variables combine to determine apparent size. Accordingly, in the present studies magnitude estimations were obtained of the apparent areas of irregular polygons; five physical areas and five viewing distances were used. In one study, high-school students served as subjects; in a second study, the same procedure was followed with college students as subjects. In scaling terms, the problem was to determine the effect of distance on the parameters of the law of apparent area. Because of its relevance to discussions of the size-distance invariance hypothesis, we used the same scaling procedure to determine the apparent magnitudes of the distances used in the main experiments.

EXPERIMENT I

Method

Materials. The stimuli were irregular polygons, constructed by Attneave and Arnoult's Method 2 (for closed-contour, angular shapes)[2] and cut from black matte cardboard. There were five different stimulus areas—8, 24, 70, 210, and 600 sq. in.—and five different polygons of each area: this made a total of 25 different figures. That each figure was irregular and unique in shape served two purposes: first, since no linear measure was common to all figures, *S* was encouraged to base his judgments on two-dimensional extent; second, *S* was not tempted to assign remembered judgments to familiar stimuli. The polygons were presented, singly, on a stand with a tripod base and a metal rod attached vertically to the base; a clip, at a height of 21 in. from the base, held the figure at its approximate center by a small metal bar cemented to its back. For all figures, the base of the stand was visible, while the portion of the rod showing ranged from 20 in., for the figure with the least extension in the vertical dimension, to 2 in., for the figure with the greatest. The stand was placed on a tabletop 30 in. from the floor.

---

[1] S. S. Stevens and M. Guirao, Subjective scaling of length and area and the matching of length to loudness and brightness, *J. exp. Psychol.*, 66, 1963, 177–186; M. Teghtsoonian, The judgment of size, this JOURNAL, 78, 1965, 392–402.

[2] F. Attneave and M. D. Arnoult, The quantitative study of shape and pattern perception, *Psychol. Bull.*, 53, 1956, 452–471.

The room used was a large classroom about 50 ft. long. The *S* sat at one end and looked out at five rows of tables, perpendicular to his line of sight, in the central portion of the room; there were also single tables, one close to him and one at the very back of the room. Some of these tables were moved to serve as convenient locations for the stand at 5, 15, 25, 35, and 45 ft. The room was not changed in any major respect; the aim was to have a normal viewing situation, in which various cues to distance would operate as they usually do.

Procedure. The instructions told *S* he was to see a series of irregularly shaped figures of different shapes and sizes, and at different distances from him; his task was to assign numbers to represent how big the figures looked to him. The *S* was encouraged to use any numbers that seemed appropriate, as long as they were proportional to the apparent sizes of the figures. There was no designated standard and no assigned modulus—that is, *S* was free to assign any number to the first figure presented.

The figures were presented one at a time: *S* turned away while *E* placed a figure, then at a signal turned back and made his judgment. Each of the five areas occurred once at each of the five distances, but which particular figure of a given area was used was determined randomly for each *S*, as was the overall order of presentation. Thus each *S* judged each of the 25 figures once. A session required about half an hour.

Subjects. There were 16 *S*s. They were recruited from summer high-school science classes and were paid for their time.

Results

The distributions of magnitude estimations for each area were positively skewed and their variances increased with their means: these are characteristics of response distributions generally found when subjects make magnitude estimations. A logarithmic transform tends to normalize the distributions and to equate the variances, and was applied to these data before an analysis of variance was performed. (There is also a theoretical rationale for such a transform: since our analysis assumed that individual subjects generate power functions, it was desirable to average data in a manner which left that relation invariant.)

All polygons of a given area were regarded as equivalent, and judgments averaged over them to yield a three-dimensional matrix (area; distance; subject). An analysis of variance indicated that the main effects, of area and of distance, on log apparent area were significant ($p < .05$), while the interaction of area and distance was not significant ($p > .05$). A subsequent analysis showed that the linear trend of log apparent-area judgments as a function of distance was significant ($p < .01$), as was the quadratic trend ($.01 < p < .05$); the variation due to higher-order trend components was not significant.

Least-squares lines fitted to log apparent area as a function of log area were obtained for each of the five distances, and the slopes, intercepts, standard errors of estimate (σ_{yx}), and coefficients of determination (r^2) are shown in Table I. The slopes are the exponents of power functions, and the intercepts are multiplicative constants; σ_{yx} and r^2 are indicators of how well the data were fitted by power functions.

The slopes, or exponents, appear to decline somewhat as distance increased. But the nonsignificant interaction between size and distance for subjects' log judgments suggests that this decline in exponents was not reliable. In contrast to the exponents, which were relatively insensitive to distance, are the intercepts, or multiplicative constants, which, as shown in Table I, increased with distance. This was indicated by the significant distance effect in the analysis of variance of subjects' log judgments: subjects tended to use larger numbers for the same stimulus area at greater distances. Table I also shows that, at all distances, over 99% of the variance in log geometric mean of apparent area (when log magnitude estimations are averaged over subjects) was accounted for by the linear regression; power functions are good fits to the averaged data.

The data are shown in Figure 1, where magnitude estimations of apparent area are plotted as a function of area, on logarithmic coordinates. Distance is the parameter, each set of symbols representing a different distance. Since the analysis of variance of subjects' log judgments showed that the interaction between area and distance was not significant, log judgments at each area were averaged over distance and a least-squares line fitted to mean log apparent area as a function of log area. The resulting line, shown in Figure 1, has a slope of .82 and represents a power function with exponent .82.

While Figure 1 represents the scaling of apparent area with distance as a parameter, the more usual form for displaying data in studies of size constancy is shown in Figure 2, where apparent area is shown, on a logarithmic scale, as a function of distance,

TABLE I

VARIOUS CHARACTERISTICS OF LEAST-SQUARES FITS OF LOG APPARENT AREA AS A FUNCTION OF LOG AREA AT FIVE DISTANCES

Distance (ft.)	Slope	Intercept	σ_{yx}	r^2
		Experiment I		
5	.866	−.896	.032	.997
15	.829	−.791	.023	.998
25	.815	−.682	.014	.999
35	.827	−.711	.045	.993
45	.786	−.638	.049	.991
		Experiment II		
5	.763	−.235	.055	.998
15	.795	−.273	.064	.998
25	.772	−.213	.090	.995
35	.771	−.213	.049	.999
45	.758	−.225	.084	.996

with area as a parameter. The same information is represented in Figure 2 as in Figure 1. The change in apparent area as area changed is reflected by the vertical placement of each curve. The influence of distance is shown by the tendency of each curve to rise as distance increased. To the extent that the five curves are

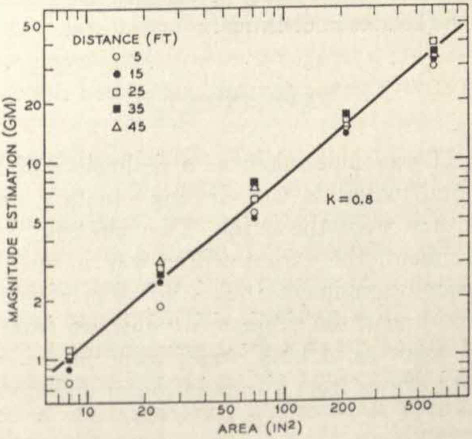


FIG. 1. Apparent area of irregular polygons as a function of physical area, with distance from the subject as a parameter; Experiment I. The magnitude estimations of apparent area are represented by geometric means. Both apparent and physical area are scaled logarithmically. The line represents a least-squares fit and has a slope of .82.

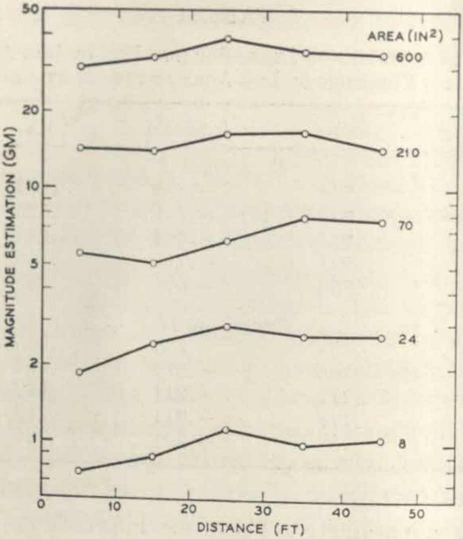


FIG. 2. Apparent area of irregular polygons as a function of distance from the subject, with area of polygon as a parameter; Experiment I. The magnitude estimations of apparent area are represented by geometric means, plotted on a logarithmic scale.

parallel, the exponent governing the growth of apparent area did not change with changes in distance.

EXPERIMENT II

Experiment II was undertaken as a replication of Experiment I. The experimental materials, the viewing situation, and the experimental procedures were the same. Two changes were made: in the first experiment, the experimenter was a research assistant, while in the second experiment, the authors served as experimenters; in the first experiment, the subjects were drawn from a population of high-school students of both sexes, while in the second experiment, they were drawn from a population of undergraduate college women.

Method

Subjects. There were 18 *Ss*, undergraduates who volunteered to participate in the experiment and who were paid for their services.

Materials. The same materials were used as in Experiment I.

Procedure. The same procedure was followed as in Experiment I. In addition, however, after each *S* had completed all 25 size judgments, she was asked to make magnitude estimations of the apparent distances of randomly selected polygons at the five distances used just before, in the main experiment: each distance was presented twice, in a different random order for each *S*. The *S* did not know before she made the size judgments that she would be asked to make distance judgments.

Results

The data of Experiment II were analyzed in the same way as were those of Experiment I. Table I gives slopes, intercepts, σ_{yx}, and r^2 for the least-squares lines fitted to log apparent-area judgments as a function of log area. Figure 3 shows the geometric means of apparent-area judgments at each of the five distances as a function of physical area; the coordinates are logarithmic. Since the analysis of variance of log judgments showed that the interaction between area and distance was not significant ($p > .05$), log judgments at each area were pooled over distances and a least-squares line fitted to mean log judgment as a function of log area. The resulting line, shown in Figure 3, has a slope of .77. This can be regarded as the mean of a distribution of slopes obtained by fitting least-squares lines to each individual's log judgments. A comparison of the mean (.77) of this distribution of individual slopes with the mean (.82) of the comparable distribution of slopes from Experiment I showed that the difference was not reliable ($t = .77$, $df = 32$, $p > .05$). In both experiments, apparent area grew as the .8 power of physical area.

Figure 4 shows apparent area on a logarithmic scale as a function of physical distance for each of the five physical areas. The analysis of variance of subjects' log judgments found the effect of distance on apparent area not significant ($p > .05$). This contrasts with the finding in Experiment I that distance did have a significant effect on apparent area. Subjects in Experiment II clearly showed size constancy—that is, their judgments of the apparent area of polygons of the same physical area showed no tendency to change, either to grow or to shrink, as distance increased.

Figure 5 shows the geometric means of apparent-distance judgments as a function of physical distance on logarithmic coordinates. The fitted line, with a slope of 1.17, accounted for over 99% of the variance in log geometric means of apparent distance; thus a

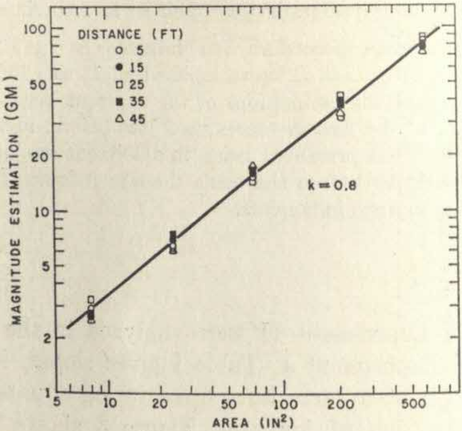


Fig. 3. Apparent area of irregular polygons as a function of physical area, with distance from the subject as a parameter; Experiment II. The magnitude estimations of apparent area are represented by geometric means. Both apparent and physical area are scaled logarithmically. The line represents a least-squares fit and has a slope of .77.

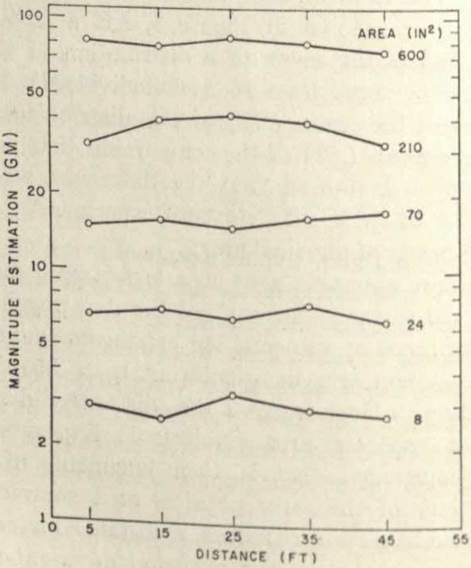


Fig. 4. Apparent area of irregular polygons as a function of distance from the subject, with area of polygon as a parameter; Experiment II. The magnitude estimations of apparent area are represented by geometric means, plotted on a logarithmic scale.

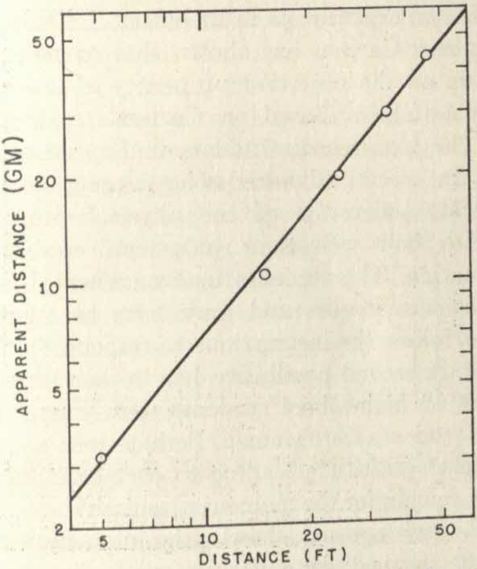


FIG. 5. Apparent distance of irregular polygons as a function of physical distance from the subject; Experiment II. The magnitude estimations of apparent distance are represented by geometric means. Both apparent and physical distance are scaled logarithmically. The line represents the least-squares fit and has a slope of 1.17.

power function with exponent 1.17 is a good fit to the apparent-distance judgments.

DISCUSSION

For all stimulus areas, apparent area did not decrease despite increases in viewing distance. As in Jenkin's earlier experiment, there was a tendency in Experiment I for the apparent area of objects of constant physical area to grow somewhat with distance.[3] It should be noted that while the increase in apparent area in Experiment I amounted to 33% from 5 ft. to 25 ft., the negative acceleration indicated by trend analysis means that this rate of increase was not maintained; indeed, apparent area at 45 ft. showed a small, though not reliable, decrease. In Experiment II, there was no tendency for apparent area to change as distance increased: the subjects assigned the same number to a near as to a far object of constant area. Why there should be a discrepancy between these two experiments in the effect of distance on apparent area is not clear. Carlson has shown that over-constancy results from emphasis on the objective properties of the stimulus situation.[4] Two possibilities, based on Carlson's finding, occur to us. One is that the high-school students in Experiment I, who were participating in special advanced-science programs, were oriented toward accurate judgments of the physical situation and were concerned with how well their judgments corresponded to the objective situation. The students in Experiment II were not predominantly science majors and may have been more willing or better able to follow the instructions to respond to the appearance of the stimuli. A second possibility has to do with sex differences. Twelve of the 16 high-school students were young men, while all of the college students were women. Perhaps men are more likely to adopt an objective attitude leading to over-constancy. These possibilities must remain for the moment speculative.

---

[3] N. Jenkin, A relationship between increments of distance and estimates of objective size, this JOURNAL, 72, 1959, 345–363.

Both experiments agree in showing no interaction between physical area and physical distance: at every distance used, ratios of apparent area increased with ratios of physical area in the same way. This invariance of the exponent of the law of apparent area with changes in distance has not previously been demonstrated, and in fact may not hold for all viewing situations, as Ross has shown.[5]

That near-constancy was obtained when a subject judged irregular and nonrepeated figures shows that while 'familiar size' may be a sufficient condition for judgments to approximate constancy, as Bolles and Bailey demonstrated, it is not a necessary condition.[6] There was no possibility in the present study of a subject's directly judging a familiar object (although many familiar objects were present). This finding agrees with that of Fillenbaum, Schiffman, and Butcher, who compared subjects' judgments of the height of a chair, a stake, and an 'abstract' figure; they found that the unfamiliar figure was judged no differently than the chair.[7]

---

[4] V. R. Carlson, Over-estimation in size-constancy judgments, this JOURNAL, 73, 1960, 199–213.

[5] H. E. Ross, Water, fog, and the size-distance invariance hypothesis, *Brit. J. Psychol.*, 58, 1967, 301–313.

[6] R. C. Bolles and D. E. Bailey, Importance of object recognition in size constancy, *J. exp. Psychol.*, 51, 1956, 222–225.

[7] S. Fillenbaum, R. Schiffman, and J. Butcher, Perception of off-size versions of a familiar object under conditions of rich information, *J. exp. Psychol.*, 69, 1965, 298–303.

The distance judgments in Experiment II indicate that in this viewing situation, apparent distance is a positively accelerated function of physical distance. The value of the exponent, 1.17, agrees closely with those reported by Künnapas (for indoor distances of about 6–65 ft.) and by Teghtsoonian and Teghtsoonian (for indoor distances of 5–45 ft. and 5–80 ft.).[8]

Of comparable importance to these particular findings is the fact that they were obtained by a technique that presents several conceptual and practical advantages. First of all, the problem of size constancy can be subsumed under the broader problem of the perception of size. Instead of investigating the effect of distance on the apparent size of a particular (and arbitrarily selected) stimulus, an experimenter can study its effect on the two parameters of the law of apparent size, $\psi = kS^n$. Distance could influence apparent size either by altering k or by altering n. In this context, it is evident that 'constancy' could characterize either of these parameters. An assumption implicit in many constancy experiments is that $n = 1$ and that constancy is therefore denoted by invariance in k. But it is at least as important to determine whether n is influenced by distance. Our result—that it is not, in at least one viewing situation—defines a previously uninvestigated aspect of size constancy.

Second, when size judgments are obtained by adjusting a comparison stimulus at a fixed location, decisions must be made about the distance of the comparison stimulus and its angular separation from the target stimuli. It is a virtue of the number-matching technique that it enables the experimenter to dispense altogether with a comparison stimulus and thus with any biasing effects associated with its location. An added advantage is that results obtained with this method are immediately and easily compared with the very many other scales of sensory magnitude obtained by number matching.

Third, and especially since the review of the literature on the size-distance invariance hypotheses by Epstein, Park, and Casey, many investigators have been convinced that in experiments in which judgments of size are to be obtained at various distances, judgments of those distances should also be gathered.[9] The difficulty in devising comparable judging techniques for measuring both apparent size and apparent distance can be overcome by the use of magnitude estimation to obtain direct judgments of both size and distance. In fact, Sternberg has shown that scaling data are uniquely well suited to evaluating several alternative versions of the size-distance invariance hypothesis.[10]

---

[8] T. Künnapas, Scales for subjective distance, *Scand. J. Psychol.*, 1, 1960, 187–192; M. Teghtsoonian and R. Teghtsoonian, Scaling apparent distance in natural indoor settings, *Psychon. Sci.*, 16, 1969, 281–283.

[9] W. Epstein, J. Park, and A. Casey, The current status of the size-distance invariance hypotheses, *Psychol. Bull.*, 58, 1961, 491–514.

Finally, the ease of our procedure should be noted. Judgments of five areas at five distances were obtained in the course of a 30-min. session. Automated stimulus presentation could halve that time. The procedure is further simplified by the absence of any equipment beyond the stimuli to be judged; the experimenter is spared the necessity of contriving some response apparatus and the subject of operating it.

In summary, the subjects matched numbers to the apparent area and distance of irregular polygons whose size and distance from the subject varied concurrently. The results were consistent with findings from size-constancy experiments which have had the subject match one size to another: apparent area remained constant (in Experiment II) or grew slightly, but reliably (in Experiment I), as distance increased. They were also consistent with experiments in the scaling of area: apparent area grew as a power function of physical area with an exponent of .8. There was no interaction between area and distance in their effects on log apparent area: the exponent .8 was the same at all distances. The exponent of the power function relating apparent and physical distance was 1.2. The direct scaling techniques (e.g., magnitude estimation) have a number of advantages for the study of size constancy: (*a*) they permit the integration of questions about size constancy into the broader framework of the perception of size; (*b*) they allow the experimenter to dispense with a variable stimulus whose definite physical properties may exert undesirable constraints on judgments; (*c*) they enable the experimenter to obtain judgments of distance as well as size in the same situation with the same scaling technique; and (*d*) they are quick and convenient.

---

[10] S. Sternberg, Four theories of the effect of distance on apparent size, paper presented at a meeting of the Psychonomic Society, 1966.

MINIMAL AVERSION THRESHOLDS FOR WHITE NOISE: ADAPTATION

RICHARD SULLIVAN, *New York University Medical Center,*
RICHARD WARREN and MARGARET DABICE,
City College, City University of New York

Abstract. Noises of four different amplitudes were presented to 20 subjects, who then judged minimal aversion thresholds for a white noise presented binaurally in ascending and descending series. Both adaptation and series influenced their judgments, but the differences in judgments were greater between series types than across adaptation levels. The findings suggest that thresholds for the attribute of aversive loudness can indeed be effectively specified. The possibility of a psychological scale of aversiveness common to all sensory systems is discussed.

Intensive sensory continua such as loudness, brightness, cold, and radiant heat all have the psychological quality of becoming aversive at subjectively high magnitudes of stimulation. Melzack and Wall have suggested that the reactive and averting aspects of pain reception depend on the exceeding of certain balanced levels of activity in both peripheral and CNS projection systems.[1] The psychological distinction between aversive and nonaversive levels of stimulation in other intensive continua may depend on a similar process, and the fact that startle, affective, and averting responses of a similar nature may be obtained by use of high-intensity stimuli of different sensory systems suggests that these continua may share common macrocerebral structures. Melzack and Casey have alluded to such a possibility in discussing the relation between affect and the sensory intensity of pain stimuli.[2] Woods and Campbell have suggested that noxious stimuli of different sensory systems may be related to a common scale of aversiveness.[3]

Received for publication by Professor E. B. Newman.

[1] R. Melzack and P. D. Wall, Pain mechanisms: A new theory, *Science,* 150, 1965, 971–978.

[2] R. Melzack and K. L. Casey, Sensory, motivational, and central control determinants of pain, in D. N. Kenshalo (ed.), *The Skin Senses,* 1968, Ch. 20.

[3] P. J. Woods and B. A. Campbell, Relative aversiveness of white noise and cold water, *J. comp. physiol. Psychol.,* 64, 1967, 493–495.

The ability of subjects to judge the aversiveness or annoyance of sound has been demonstrated in a number of studies.[4] So has the ability of subjects to judge the discomfort of and minimal aversion thresholds for brightness.[5] And behavioral indexes of the relative aversiveness of sensory stimuli have been obtained by cross-modal comparisons of noise with electric shock and with cold water.[6] Undoubtedly, the attribute of sensory aversiveness is due to a complex interaction of peripheral and central factors similar to that suggested by Kryter for the attribute of noisiness.[7] However, Bourassa and Wirtschafter (see n. 5) reported that while other factors may have contributed to the judgment of visual discomfort, the major variable appeared to be total luminous flux. They speculated that the process responsible for the discrimination of discomforting levels of visual stimuli may be the same as that suggested by Melzack and Wall (n. 1) for the appreciation of pain stimuli. It is possible that the transition from nonaversive to aversive levels of stimulation may be related to a common process across sensory modes.

The ability of subjects to judge aversiveness of stimuli suggests that a transition point, or minimal aversion threshold (MAT), could be specified which would divide the loudness continuum into ranges of aversive and nonaversive stimuli.[8] As Kryter (nn. 4, 7) has noted, the band of noise used, the adaptation level, and the subject's previous experience with noise influence judgment of such an attribute. For example, Spieth (n. 4) found that individuals who judged the annoyance of various bands of white noise on the basis of their own work environments had different thresholds. However, both Kryter (n. 7) and Spieth (n. 4) found that the range of annoyance judgments was relatively restricted. If the specification of an effective minimal aversion threshold is to be meaningful, the range of values obtained for different adaptation levels must also be limited. That is, while adaptation level may influence judgments of the threshold, there must be a point at which adaptation level has a diminishing affect on the discrimination of a threshold of minimal aversiveness. The present experiment was thus designed to determine what effect different levels of adaptation have on the judgment of a minimal aversion threshold for white noise.

---

[4] D. A. Laird and K. Coyle, Psychological measurements of annoyance as related to pitch and loudness, *J. acoust. Soc. Amer.*, 1, 1929, 158–159; W. Spieth, Annoyance threshold judgments of bands of white noise, *J. acoust. Soc. Amer.*, 28, 1956, 872–877; K. D. Kryter, Psychological reactions to aircraft noise, *Science*, 151, 1966, 1346–1355; R. Sullivan, Subjective matching of anxiety to intensities of white noise, *J. abnorm. Psychol.*, 74, 1968, 646–650.

[5] C. M. Bourassa and J. D. Wirtschafter, Mechanism of binocular increase of discomfort to high luminance, *Nature*, 212, 1966, 1503–1504; R. Sullivan, Minimal aversion thresholds for brightness under light and dark adaptation conditions, this JOURNAL, 81, 1968, 581–585.

[6] B. A. Campbell and J. M. Bloom, Relative aversiveness of noise and shock, *J. comp. physiol. Psychol.*, 60, 1965, 440–442; Woods and Campbell (n. 3 above).

[7] K. D. Kryter, Concepts of perceived noisiness, their implementation and application, *J. acoust. Soc. Amer.*, 43, 1968, 344–361.

[8] Sullivan (nn. 4, 5 above).

METHOD

Subjects. The *S*s were 20 undergraduates (age range, 18–22) with normal hearing.

Apparatus. The apparatus consisted of a white-noise generator in circuit with an amplifier, a series of decade resistors, and a set of calibrated earphones. Calibration of noise amplitudes was done through use of a Ballantine VTVM. The scale used by *E*s to set levels of stimuli and to record judgments was independently assigned an arbitrary code, and the two *E*s who ran *S*s were unaware of the amplitudes of noise presented during the various trials. The noise was presented binaurally while *S*s were seated in a darkened, sound-attenuated cubicle.

Procedure. The *S*s were instructed that the experiment dealt with their "ability to detect a level of noise that indicated a transition from nonaversive sound to aversive sound." They were told repeatedly throughout the testing periods that they were to concentrate on the discrimination of a sound which was minimally aversive, annoying, and bothersome. All *S*s were given a number of practice trials with the test stimuli being presented for both ascending- and descending-noise series. They were then told that the test noise would be preceded by a brief noise (adaptation stimulus), which would remain on at a certain level and then be turned off. Immediately after the turning off of the adaptation stimulus, a noise would be presented that either increased or decreased in amplitude; they were to judge this second noise. For the ascending-noise series, *S*s were told to say 'Now' into a microphone when the noise level was at the point of becoming aversive, annoying, and bothersome. For the descending-noise series, they were told to say 'Now' when the noise level was at the point of no longer being aversive and annoying.

The adaptation-noise levels used were 70, 85, 95 and 110 db. (±1 db., re .0002 dynes/cm^2). The range of noise used for the test series was either 50–95 db. or 50–110 db.[9] The manner of presentation of both adapting and test stimuli was as follows. The adaptation-noise level was attained by turning up the amplitude from a standard level of 50 db. at different rates of increment for different adapting stimuli (Table I). When the proper adaptation level was reached, the noise was left on for varying intervals depending upon the amplitude of the adapting stimulus (Table I). Following the turning off of the adaptation noise in the *ascending* trials, white noise was increased from the standard 50-db. level over 60-, 90-, and 120-sec. intervals to a level just beyond that considered by the *S* to be the minimally aversive level, and then it was turned off. A rest period of 2 min. followed each trial. In the *descending* series, following the turning off of the 70- or 85-db. adapting stimulus, the test noise was turned on at 50 db., rose to 95 db. in a 2-sec. interval, and then decreased to a point just beyond that considered by the *S* to be the minimally aversive level. The same procedure was followed for the 95- and 110-db. adaptation levels, with a rise time of 3 sec. to the 110-db. level. Thus, the upper anchors in the descending series differed, from 95 db. for the 70- and 85-db. adaptation levels, to 110 db. for the 95- and 110-db. adaptation levels.[10] The procedure is outlined in Table I.

---

[9] The upper noise amplitude of 110 db. seemed sufficient to demonstrate adaptation effects while remaining within tolerable psychological and physiological limits.

Using a mixed-models design,[11] *S*s were given two ascending and two descending trials for each adaptation level. Times of ascent and descent were balanced, then randomized over trials. All *S*s were tested over a period of a month with an interval of approximately a week intervening between sessions.

TABLE I

ADAPTATION AND TEST PROCEDURES

Adaptation stimuli			Test stimuli				
			Ascending series		Descending series		
Level (db.)	Rise time (sec.)	Adaptation time (sec.)	Noise range (db.-MAT)	Ascent rate (db/sec)	Rise time (sec.)	Noise range (db.-MAT)	Descent rate (db/sec)
70	2	12	50→MAT	.5–2	2→	95→MAT	.5–2
85	3	8	50→MAT	.5–2	2→	95→MAT	.5–2
95	4	6	50→MAT	.5–2	3→	110→MAT	.5–2
110	6	4	50→MAT	.5–2	3→	110→MAT	.5–2

Note: All db. values ± 1 db. re .0002 dynes/cm^2; ascent and descent rates are approximate figures.

RESULTS

Means, standard deviations, and coefficients of variation for ascending and descending judgments of the minimal aversion threshold for white noise at the four levels of adaptation are presented in Table II. Mean differences between ascending and descending judgments for each level of adaptation were greater than mean differences across adaptation levels within the ascending or descending series. Both absolute and relative measures of variability of judgments suggest a substantial series effect in the judgments of threshold.

---

[10] The use of the different upper anchors, 95 and 110 db., for the descending trials was necessitated by the possibility that if 110 db. were used for the 70- and 85-db adaptation series it might have considerably altered the adaptive state of the system.

[11] Q. McNemar, *Psychological Statistics*, 1962.

Table III presents a three-way analysis of the variance of the judgments (four adaptation levels; ascending versus descending series; subjects). Difference in the judgments of the threshold were significant for each of the four adaptation levels [*F* (3, 57)

TABLE II

MEANS, STANDARD DEVIATIONS, AND COEFFICIENTS OF VARIATION OF JUDGMENTS AT FOUR LEVELS OF ADAPTATION

Series	Adaption level			
	70 db.	85 db.	95 db.	110 db.†
Ascending				
Mean judgments	79.9	80.2	84.9	84.2
Standard deviations	10.3	9.2	8.3	8.2
Coefficients of variation	12.9	11.5	9.8	9.7
Descending				
Mean judgments	87.5	88.8	93.8	92.6
Standard deviations	4.3	3.9	5.2	6.1
Coefficients of variation	4.9	4.4	5.5	6.6

Note: *N* (*S*s) = 20; total judgments = 320.
† All values ± 1 db. re .0002 dynes/cm^2.

TABLE III

ANALYSIS OF VARIANCE

Source	*df*	*MS*	*F*
Subjects (A)	19	249.3	
Series (B)	1	2,822.4	31.6***
Adaptation level (C)	3	317.4	18.6***
A × B	19	89.2	6.7**
A × C	57	17.1	1.2
B × C	3	3.2	.2
A × B × C	57	13.3†	
Total	159		

† *MS* error term.
*** *p* .001.
** *p* .005.

$= 18.6, p < .001$]. Differences between the ascending and descending series were also significant [$F\ (1, 19) = 31.6, p < .001$].

Matched-pair tests of the difference between mean judgments of the threshold at the four adaptation levels for both ascending and descending series indicated insignificant differences for the 70–85-db. and 95–110-db. comparisons. The 85–95-db. comparison was significant for both ascending and descending trials [ascending, $t\ (19) = 4.12, p < .001$; descending, $t\ (19) = 9.61, p < .001$].

DISCUSSION

The results indicate that different adaptation levels significantly influenced judgment of the minimal aversion threshold for white noise. However, the effect was not an increasing monotonic function of the adaptation level. Thus, while a difference of 10 db. in adaptation level (85–95-db. comparison) significantly altered judgments in both ascending and descending trials, differences of 15 db. (70–85-db. and 95–110-db. comparisons) did not alter the judgments. In terms of the previously reported mean annoyance or aversion threshold of 85 db.,[12] and of the overall mean in the present study of 86.5 db., the significant effect of the adaptation noise may have been dependent on whether the noise was above or below a subjective level of aversiveness. It may be that the *relative* effect of adaptation on judgment of a threshold of noise aversiveness diminishes as adaptation levels increase or decrease from the intensity range of 85–95 db. Furthermore, the effect of adaptation on judgment of the threshold seems to have been less than the effect of series. For example, the maximum mean difference between adaptation levels (6.3 db., 70–95-db. comparison, descending series) was less than the maximum mean difference between ascending and descending judgments within an adaptation level (8.9 db., 95-db. adaptation level). The coefficients of variation also indicated that the series effect was greater than the adaptation effect. The series effect was independent of the adaptation effect (Table III; B × C interaction term insignificant). Consequently, the present findings suggest that a minimal aversion threshold for white noise could effectively be specified for rather broad ranges of adapting noises. However, the results further indicate that such a specification might not be reliable or valid if ascending and descending judgments of MAT noise were combined and averaged.

[12] Spieth (n. 4); Sullivan (n. 4).

The significant effect of series may have been due to the anchor stimuli used in the present study (50 and 110 db.) or to subjective attitudes toward the judgment. Since the lower and upper anchor values were the same for each adaptation level, it would appear that while anchoring may have contributed to the significant series effect, its influence was not as great as that suggested in the difference between ascending and descending judgments. It seems likely, in terms of the coefficients of variation, that the series effect derived from the ease with which the subjects could demarcate the aversive range from the nonaversive range of noise stimuli. That is, when a subject is presented with a nonaversive noise that is gradually increasing in amplitude, he must, in a sense, anticipate when the noise will become minimally aversive or annoying. However, when the level of noise is already within an aversive range and is gradually diminishing in amplitude, the judgment of the absence of the attribute of aversiveness is more readily accomplished. Subjects in general have had more experience avoiding noise sources that are aversive and overload the system. They would appear to be better discriminators of when the overload has been removed than of when the overload is beginning to approach some upper limit of the auditory system.

The subjective appreciation of an overloading of the auditory system may parallel physiological changes occurring due to alterations in the mechanics of the middle ear, impulse integration in the cochlea, or central factors. Thus, mechanically, the two middle-ear muscles that pull the drum and stapes, and permit accommodation of wide ranges of amplitude, could be in part responsible for the judgment of auditory aversiveness. These middle-ear muscles are analogous to the muscles that control the iris.[13] It may be that Bourassa and Wirtschafer's finding (see n. 5) that muscles of the iris were not entirely responsible for the judgment of visual (brightness) discomfort could be applicable to the contribution of the middle-ear muscles to the judgment of auditory annoyance.

The judgment of auditory aversiveness occurs well below the amplitude required to obtain a shift in the axis of rotation of the stapes.[14] This suggests that the major factor in the appreciation of aversiveness of sound within the range of approximately 85–120 db. could be neurological. Noises below this level would be judged aversive if the adaptation-noise level was low and the test noise was presented suddenly (impact stimulus). The aversiveness of the lower noise amplitudes would depend on the relative increase in amount of neuronal activity. The aversiveness of the higher amplitudes would depend on both the nature of stimulus presentation and the relative ability of the system to accommodate such amplitudes. Thus, while lower amplitudes than 85 db. could be experienced as aversive, adaptation to the noise would be accomplished within a short period of time, and the noise would no longer be experienced as aversive. With the high amplitudes the adaptive capacity of the auditory system would be limited and the aversiveness of the noise would be experienced throughout its duration. Below the level of approximately 120 db., 'cerebral' factors (e.g., a job in a very noisy factory) could inhibit the reaction to the noise. Above that level, however, inhibition of averting responses would become increasingly difficult.

---

[13] G. von Békésy and W. A. Rosenblith, The mechanical properties of the ear, in S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951, 1085.

[14] G. von Békésy, *Experiments in Hearing*, 1960, 112–113.

The action system responsible for the averting reactions in response to pain stimuli suggested by Melzack and Wall (n. 1) may also be responsible for the averting reactions to noise stimuli. That is, while these systems would have independent projection areas, they might share with other intensive sensory systems a common central structure responsible for the startle and averting responses as well as the accompanying affect states. Thus, whenever there was a sudden increase in the relative level of activity in any of these systems, or when the stimulation approached the upper load limit of the system, the balance of activity in the action system would be altered and an averting response would occur.

If a common psychological scale of aversiveness were to be found for differing sensory systems, it would lend support to the hypothesis that levels of activity in macrocerebral structures are related across sensory modes, and it would supplement the evidence, from Green and Arduini's recordings from hippocampal regions, for example, that different sensory systems seem to share certain general effects.[15] The attempt to relate physiological measures to psychophysical judgment[16] would also be more readily accomplished if a common scale were shared by different intensive modalities.

---

[15] J. D. Green and A. A. Green, Hippocampal electrical activity in arousal, *J. Neurophysiol.*, 17, 1954, 533–557.

[16] H. D. Kimmel, Equating stimulus intensities by means of the GSR, *Psychon. Sci.*, 7, 1967, 77–78.

TEMPORAL SHIFTS IN THE VERBAL BEHAVIOR OF INTRODUCTORY PSYCHOLOGY STUDENTS

Thomas W. Turnage, *University of Maryland*

Abstract. A series of three studies using association and paired-associate learning tasks with course-related materials showed that both changes in 'verbal responsiveness' and opportunities for associative interference were occurring over the semester in the population sampled. Some general implications of these findings for studies using similar pools of subjects and for studies of retroactive mechanisms in long-term retention are briefly considered.

The classroom situation is one in which the student is systematically exposed to course-related material in a more or less 'natural' linguistic environment, and where he is periodically examined for behavioral changes that reflect learning and retention of such items. Thus, the classroom can be compared to the laboratory situation in which a subject is also required to learn and recall materials to which he has been systematically exposed; even though the linguistic situation in the laboratory is somewhat more 'artificial' and the practice sessions are shorter. This comparison suggests that the use of course-related materials as stimuli and of students in the course as subjects provides a means of examining two topics of general importance to researchers in verbal behavior. The first of these relates to artifacts that may be introduced into the experimental situation when subjects, especially those from the introductory course in psychology, are sampled at different intervals in

Received for publication June 8, 1970. The research was supported in part by a General Research Board Grant from the University of Maryland and in part by a grant from the Bio-Medical Research Support Committee to the Center for Language and Cognition. Computer time was supported by National Aeronautics and Space Administration Grant NS-398 to the Computer Science Center at the University of Maryland. The author thanks a number of graduate and undergraduate students who assisted in data collection and analysis, particularly Robert Voytas, Joe Jaffa, Meredith Friedman, Carolyn Crane, and Forest Werner.

the term;[1] and the second concerns the role of relatively long-term *retroactive* mechanisms, as opposed to proactive ones, that may act on verbal materials with different probabilities of occurrence in the subject's linguistic environment.[2] These considerations provided the major impetus for the three studies reported in this paper.

METHOD

Subjects. Over a period of three semesters, a total of 592 *S*s, fulfilling a requirement of an introductory course in psychology at the University of Maryland, participated in the experiments. Of these 592, 314 *S*s were used in Experiment I, 158 were used in Experiment II, and 120 in Experiment III.

Materials for Experiments I and II. The general pool of experimental materials was composed of psychological terms, which were selected from the vocabulary sections of textbooks for the introductory course in psychology, and control words. The control words did not appear in the vocabulary sections, but they matched the psychological words on the basis of part of speech, number of syllables, initial letter, and 'L'-count values.[3] This gave a pool of approximately 150 words. Each of these words could be characterized as either high-frequency psychological, high-frequency control, low-frequency psychological, or low-frequency control. The median 'L'-count values for the low-frequency psychological and control words were 2.9 and 3.2, while for the high-frequency psychological and control words they were 783 and 804. Eight-page booklets containing two different high-frequency psychological and two different low-frequency psychological words and their appropriate control words—with one word per page, and with pages randomly ordered—were presented to *S*s for free-association tests, with 60 sec. allowed for associating to each word on a page.[4] There were 16 different exemplars of the basic booklet, involving 128 different words, and an attempt was made to use these different booklets equally often at each time of testing.

Procedure for Experiment I. At the beginning of the semester, *S*s were tested in their discussion sections by their discussion leaders, with 8 discussion sections being used. At semester's end—three months later—11 *different* sections were tested in the same manner to secure an additional 157 *S*s, the *N* for the first test.

Procedure for Experiment II. The same procedure was followed as in Experiment I except for the variation that the *same* *S*s were tested with different stimuli at the end of the semester. Thus, since a *different* booklet was used by a given *S* for the two tests, his second association task was not simply a 'retention' task. Of the 158 *Ss* used in the first test, 38 were unavailable for the second test.

---

[1] B. J. Underwood, E. Schwenn, and G. Keppel, Verbal learning as related to point of time in the school term, *J. verb. Learn. verb. Behav.*, 3, 1964, 222–225.

[2] B. J. Underwood and L. Postman, Extra-experimental sources of interference in forgetting, *Psychol. Rev.*, 67, 1960, 73–95.

[3] E. L. Thorndike and I. Lorge, *The Teacher's Word of Book of 30,000 Words*, 1944.

[4] For general method and instructions, see C. E. Noble, An analysis of meaning, *Psychol. Rev.*, 59, 1952, 421–430.

Procedure and materials for Experiment III. The purpose of this experiment was to magnify potential interference effects over time by using a paired-associate learning task involving an A-B, A-Br design. Two judges selected 20 pairs of 'appropriate' associates (A-B pairs) from the pool of psychological words (e.g., *validity-reliability, paranoia-psychosis*). Median 'L'-count values for A and B items respectively were 0 and 56. These A-B pairs were then divided into two lists of 10 pairs. The resulting lists of paired associates were judged to be comparable in terms of type of pairs (e.g., 'clinical,' 'statistical,' etc.). Two corresponding lists of A-Br pairs (e.g., *validity-psychosis, paranoia-reliability*) were formed by re-pairing the response terms of the A-B lists. Unfortunately, due to a procedural error, only 7 of the 10 A-B responses were used in the re-pairing. However, the remaining 3 responses on each of the A-Br lists were judged 'inappropriate' for the stimuli with which they were paired and, therefore, did not seriously affect the experimental design. Three random orders were assigned to the pairs for the four experimental lists.

The pairs from a given list were presented by means of a slide projector at a 2-sec. rate. Then, the stimulus terms were presented alone at a 3-sec. rate (study-test method), during which time the *Ss* wrote the responses. There was a 10-sec. intertrial interval between trials, and each *S* had six acquisition trials on the same list.

There were 80 *Ss* tested for the first time at the beginning of the semester, and 40 different *Ss* tested for the first time at the end of the semester. The *Ss* were tested in groups of four to six, and the groups were assigned to experimental conditions, on the basis of random numbers, as they entered the laboratory. Of the 80 *Ss* tested at the semester's start, 59 (74%) returned for a second test, which was a relearning task. The 21 *Ss* who failed to return were from groups scattered across all experimental conditions.

RESULTS

Changes in m *over time.* Table I shows the mean number of associations, or m, as a function of the various experimental conditions in Experiments I and II. Note that m generally showed an absolute increase over time for all types of words and that the usual direct relationship between frequency and m obtained (see n. 4).

Two separate analyses of variances for the experimental results shown in Table I indicated that time was a reliable source of variation in Experiment I [F (1, 312) $= 6.21, p < .05$] and in Experiment II [F (1,119) $= 11.40, p < .01$] but did not interact with type of word (psychological versus control) nor with frequency in either of these experiments ($Fs < 1.08$). Frequency, of course, was a significant source of variation in both Experiment I [F (1,312 $= 590.37$, $p < .01$] and Experiment II [F (1,119) $= 344.60, p < .01$].

TABLE I

MEAN NUMBER OF ASSOCIATIONS PER ITEM AS A FUNCTION OF EXPERIMENTAL CONDITIONS: EXPERIMENTS I AND II

Time of test	High-frequency words Psych.	High-frequency words Control	Low-frequency words Psych.	Low-frequency words Control
		Experiment I		
Beginning of semester	8.92	9.62	6.97	6.92
End of semester	9.46	10.49	7.68	7.69
		Experiment II		
Beginning of semester	9.04	9.36	7.17	6.27
	(9.04)	(9.95)	(7.22)	(6.49)
End of semester	9.30	9.81	7.80	8.86

Note: Means in parentheses are for subjects who were not available for testing at the end of the semester.

Paired-associate learning. In order to attenuate any effects of subject selection over the semester, the following procedure was followed before the initial data analysis for Experiment III. Of the 59 subjects who returned at semester's end for their relearning task, 40 were selected to be compared (on their initial test) with the 40 subjects tested at semester's end. Selection from among the 59 was by random numbers with the restriction that there be 10 subjects for each of the four experimental lists used. Table II shows the mean performance for the resulting pool of 80 subjects as a function of the basic experimental conditions.

TABLE II

MEAN NUMBER OF PAIRS CORRECT PER TRIAL AS A FUNCTION OF EXPERIMENTAL CONDITIONS: EXPERIMENT III

	Time of test			
	Beginning of semester		End of semester	
Trial	A-B lists	A-Br lists	A-B lists	A-Br lists
1	4.75	1.80	4.00	0.80
2	6.30	3.70	5.75	2.70
3	7.80	5.25	7.35	3.75
4	8.75	5.85	8.15	4.50
5	9.20	6.50	8.50	5.20
6	9.05	7.50	8.60	6.05
Col. means	7.64	5.10	7.06	3.83

An analysis of variance for the data in Table II indicated a reliable *drop* in the paired-associate learning rate over the semester, with the mean per trial being 6.37 at the beginning of the semester and 5.44 at the end of the semester [$F\ (1, 76) = 4.78, p < .05$]. In general, A-B lists were learned faster than A-Br lists; with the respective means being 7.35 versus 4.47 [$F\ (1, 76) = 45.46, p < .01$]. Thus, even early in the semester, the subjects showed difficulty learning 'inappropriate' pairs. The effect of trials was also significant [$F\ (5, 380) = 155.19, p < .01$], as would be expected. However, none of the remaining sources of variation was reliable ($Fs <$ 2.00), although the interaction of time and lists was in the expected direction (see Table II).

The relearning data for subjects returning from the first test were not of direct interest in the present study and are not reported in detail. However, it may be noted that there was evidence for reliable drops in retention performance over the semester—with A-Br performance on the first relearning trial being poorer than that for A-B lists—and for substantial savings in relearning. The more rapid forgetting found for the A-Br lists is, of course, difficult to interpret in light of the differences in original learning that also obtained (see Table II).

DISCUSSION

Verbal responsiveness. It seems quite clear that changes which might be described generally as increases in 'verbal responsiveness' occurred in the subjects over the course of the semesters involved in Experiments I and II. However, the precise causes of these changes are far from apparent. One possibility to be considered is that the temporal shifts in m values reflect the 'well-known' fact that subjects who are 'available' at semester's end have been selected on the basis of ability, and that this is correlated with performance in tasks like those used in the present studies. Be that as it may, such an explanation does not seem adequate for the present data for two reasons. First, more than 75% of the original sample was retested in Experiment II, thus minimizing factors of subject selection; yet, changes in m values in this study were comparable to those found in Experiment I, where subject selection might have been a more important factor. Second, those 38 subjects who were not available for retesting in Experiment II (see Table I) tended toward higher m scores; thus, it seems likely that the increases in m

noted would have been even greater had these 38 subjects been 'recaptured.' (It may also be noted that an unpublished study in this series, not included in this report, indicated general increases in m when 95% of the original sample of subjects was retested.)

Correspondingly, however, one must consider the possibility that the temporal shifts in Experiment II reflected either (*a*) specific 'skills,' learned during the first test, that persisted over three months or (*b*) general transfer effects resulting from the fact that the subjects had previously participated in this experiment, and possibly 'outside' ones, because of the requirements of the course. These alternatives also fail to be convincing.

The specificity explanation could not obtain, of course, for Experiment I, and it does not seem likely for Experiment II because of the similar trends over the three-month period for both experiments and the long intertest interval involved. The general transfer explanation is not as easily dealt with on the basis of the present data. However, this explanation also seems unlikely when we consider the finding of Underwood et al. (n. 1) that performance on a paired-associate task was virtually identical for subjects tested during the first half of the term, when they had no prior laboratory experience, and for subjects tested during the second half of the term, when they had one or two previous laboratory experiences (not verbal-learning experiments). Of course, there are important differences between the present methodology and that of Underwood et al., which must be considered. Nevertheless, general transfer effects do not appear to be critical factors in the present experiments.

Temporal changes in paired-associate learning. The next point of discussion is the discrepancy between present learning trends over time for the paired-associate data and the negative findings for such trends reported by Underwood et al. (n. 1). The most obvious reason for the discrepancy is that the present study used experimental materials especially selected to make it likely that the subjects would encounter them in learning situations over the course of the semester. As noted before, the purpose of such selection was to investigate the feasibility of detecting *retroactive* effects of linguistic experiences on verbal behavior, by using course-related material of a highly specific nature. If such a technique is fruitful, it may be quite useful for theoretical purposes, since logical considerations and certain inherent limitations of laboratory tasks have led to a contemporary focus on *proactive* transfer in the analysis of forgetting (see n. 2). Happily, Experiment III offered some evidence for the success of such a strategy, as the data in Table II indicate. Note especially the extremely poor performance of the A-Br groups at semester's end.

As were the *m* data, the temporal changes in paired-associate learning are difficult to explain by simply assuming that subject-selection factors were involved. This 'selection' explanation seems especially unlikely for the data on the learning of paired associates, since (*a*) all comparisons involved only subjects who were available for testing at the end of the semester and (*b*) the direction of the learning differences over time was the reverse of what would be expected if superior learners were being 'selected for' over the course of the semester. Considering the report of Underwood et al. (n. 1), it also seems unlikely that 'motivational' changes in the subjects who were available for experiments at semester's end could explain the present data. Therefore, alternative explanations must be considered. A likely one is that the observed changes in the learning of paired associates reflected—at least in part—associative interference deriving from the subject's increased 'verbal responsiveness' (increased vocabulary?).

Temporal changes in general versus technical associations. In an attempt to isolate specific verbal mechanisms that might be responsible for deterioration in A-B and A-Br learning over the semester, the associations from Experiment I were selected for further analysis. It was hoped that these data would be minimally influenced by a set on the part of the subjects to give particular classes of associations or to recall particular examples of associations as a function of prior testing. Two judges sorted the associations into two categories: namely, (*a*) psychological associates and (*b*) nonpsychological associates. The means resulting from this analysis are shown in Table III as a function of type of stimulus word (psychological versus control) and time of association test.

The means in Table III show that psychological stimuli elicited more psychological associates than did the control stimuli, which was to be expected; and that the bulk of the associates to both psychological and control stimuli were of the nonpsychological variety. However, these nonpsychological associates were generally quite 'relevant' ones (e.g., *money, award, deed,* and *wanted* were associates to *reward*), and they did not seem to be the result of extensive 'chaining' of associates (n. 4) by the subjects. Of greater

TABLE III

MEAN NUMBER OF PSYCHOLOGICAL AND NONPSYCHOLOGICAL ASSOCIATES AS A FUNCTION OF TIME AND TYPE OF STIMULUS WORD: EXPERIMENT I

	Psych. stimulus words		Control stimulus words	
Time of test	Psych. associate	Nonpsych. associate	Psych. associate	Nonpsych. associate
Beginning of semester	2.56	5.38	1.69	6.58
End of semester	2.57	6.00	1.88	7.21

importance is the finding that nonpsychological associates increased over time for both psychological and control stimulus words, as if reflecting a general change in verbal responsiveness not directly related to course materials. Although not expected, this outcome was consistent with the finding (Experiment III) that paired-associate learning was generally depressed over time when psychological words were used as the stimuli. The implication, of course, is that the number of nonpsychological associates increases over time and hence provides sources of associative interference in situations like the one being considered.

Although the data in Table III flatly deny that there was any sizable increase in the *number* of psychological associates to psychological stimuli over time, they do provide indirect evidence for an increase in the *strength* of such associations when considered with the paired-associate data in Table II. That is, the constant superiority of A-B versus A-Br learning over the semester—with the suggestion of a relative increase by semester's end—may mean that A-B strength was increasing, and not just 'holding its own' in the face of interference from nonpsychological associates. Such an interpretation is at least not improbable, and it may be reassuring for many course instructors.[5]

In conclusion, the subject population sampled in the three experiments appeared to be 'different,' or at least more 'verbally responsive,' at semester's end. Whether such 'differences' might in-

[5] With respect to this point, it should also be noted that exploratory studies by Verplanck and his associates at the University of Tennessee indicate that associations to course-related materials may be quite useful for assessing academic success and grasp of concepts in the classroom (W. S. Verplanck, personal communication, 1968).

troduce artifacts into experiments that use similar subjects sampled at different points in time would seem to depend on the nature of the experimental problem, materials, and method. Of greater potential importance for our present purposes is the fact that such temporal 'differences' in verbal repertories might be used to good theoretical advantage in examining the role of intervening linguistic activities on previous or subsequent verbal-learning tasks.

NOTES AND DISCUSSION

THE SIXTY-SIXTH ANNUAL MEETING OF THE SOCIETY OF EXPERIMENTAL PSYCHOLOGISTS

The sixty-sixth annual meeting of the Society of Experimental Psychologists was held at the University of Pennsylvania, Philadelphia, on April 10 and 11, 1970. Richard Solomon, Chairman of the Society for the year, presided at the business meeting and at sessions for the presentation of scientific papers.

Thirty-eight members and fellows were present: Amsel, Blough, Cofer, Eriksen, Estes, Geldard, E. G. Gibson, J. J. Gibson, Grant, Green, Grice, Hess, Hirsh, Hochberg, Humphreys, Hurvich, Irwin, Jenkins, Kamin, Kappauf, Kendler, Kessen, Kimble, Lindsley, Luce, Mandler, Meyer, Morgan, Pratt, Ratliff, Riggs, Solomon, Stellar, Teitelbaum, Thomas, Underwood, Volkmann, and Wickens.

Dorothea Jameson Hurvich, Alvin M. Liberman, Ulric Neisser, and Karl Pribram were elected to membership in the Society, bringing the total to 73 members and 41 fellows.

Reports of research were presented in scientific sessions on both days of the meeting. The Warren Medal for 1970 was awarded to Daniel Lehrman "for his impressive program of research showing the interaction between hormonal and behavioral factors in the timing of a series of courtship, reproductive, and infant care activities in the ring dove."

The Society accepted the invitation of the University of Colorado to meet in Boulder, Colorado, in 1971. Gregory Kimble was elected Chairman for 1970–1971. Carl Duncan was elected Secretary-Treasurer.

William E. Kappauf, secretary;
University of Illinois at Urbana-Champaign

BOOK REVIEWS

Edited by JULIAN HOCHBERG,
Columbia University

SELECTIVE ATTENTION AND EVOKED POTENTIALS. By RISTO NÄÄTÄNEN. Helsinki: Suomalainen Tiedeakatemia, 1967. Pp. 226.

The monograph is not so important a form of scientific reporting as it once was. It is true that monographs do appear regularly, but these are usually subject to the same constraints as are journal articles. They are just larger. Their authors may not be too speculative. They must not take too much time to tell us about the history of their subject. And they often use statistical tests as substitutes for closely reasoned and satisfying treatments of the processes that might underly the phenomena forming the subject matter of the monograph. Of course, there are notable exceptions to this trend. Näätänen's monograph is close to being one of these exceptions.

Näätänen takes the time to develop fully the reasoning underlying his experiments and those of previous authors. He is also willing to speculate about the implications of his results and their relationship to the results of others. He tries to be as clear as possible about what he did and why he did it. All of this makes for interesting reading. The failings of the monograph do not stem from Näätänen's failure to make full use of the medium. They are the failings from which we all suffer when we discuss the elusive problem of attention.

The heart of the monograph is a set of three experiments done while Näätänen was at Lindsley's laboratory. These experiments were designed to provide a more adequate test of the notion that the evoked potential changes with selective attention. Stated baldly, it has been held that attention reflects the operation of a filter. If something is attended to, then signals arising from that stimulus are enhanced when they reach the cortex. If the stimulus should be presented while the subject is doing something else, then the signals it produces are relatively attenuated. This is the well-known interpretation of Hernández-Peón, and of many other investigators of the neural processes that underlie attention. The discovery by Dawson that it is possible to detect a stimulus-connected electrical response at the human scalp led to a spurt of enthusiastic attempts to show that analogues to the results that are obtained in experiments with animals may also be obtained with human subjects.

The results obtained with animals are typified by the experiment of Hernández-Peón in which potentials that are evoked in the cochlear nucleus of the cat, by clicks, are attenuated when a mouse is placed in the cat's field of view. This was construed to be an instance of selective attention, inasmuch as the cat's attention to the click was distracted by the presence of the mouse. It was assumed that the neural activity produced by the now-irrelevant stimulus (the click) was blocked at the receptor. Näätänen tells the story of such experiments very well. He also tells us about the devastating criticisms of these pioneering experiments—particularly those raised by Horn and his colleagues. Näätänen adds some of his own criticisms and makes an entirely convincing case that such results do not arise from the action of selective attention. He mentions many possible causes for such effects. One of them is the visual search behavior of the animal when a novel stimulus is introduced. Another is the possibility that the ascending reticular activating system simply changes the state of the animal's brain and causes the evoked response to diminish. He argues that the only proper experiment would be one in which the brain response to the novel stimulus is measured while the response to the click is measured. Selective attention should cause the response to a stimulus to be higher when the stimulus is novel than when the same stimulus is no longer novel. Also, the response to the ongoing and irrelevant stimulus should be depressed. Näätänen cites experiments which suggest that this is not the case. During arousal by some novel stimulus, all evoked responses are altered in the same way. Usually these responses in animals are diminished by the introduction of novelty, but sometimes they are uniformly elevated. All in all, the results of animal experiments have failed to give us conclusive evidence for the presence of a selective filter. Näätänen reasons that perhaps it is really too hard to tell if a cat is or is not paying attention. The human could be a better 'preparation.'

In humans, the measurement of the evoked potential requires the use of signal averaging. Responses evoked by a stimulus are hidden in the background EEG activity. They cannot be seen in any detail when recorded one response at a time. One way around this problem is to superimpose the epochs of EEG activity that follow immediately after each stimulus presentation. If, as assumed, the response has a largely constant phase lag after the stimulus, it would tend to add up as succeeding epochs of EEG activity are added together. The background EEG activity, however, does not necessarily have a constant phase relationship to the stimulus. Hence, from one epoch to the next the succeeding trains of EEG activity would change phase and tend to become self-canceling. In this way the unwanted signals comprising the EEG (the 'noise') are attenuated relative to the wanted signal of the evoked potential. This should allow us to get a picture of the average, and hopefully typical, response of the organism to the stimulus, if we stimulate often enough.

In a typical experiment, the subject would have to detect an occasional dim flash of light in a train of otherwise bright flashes. It is obvious that if he does this, then he is paying attention to the flashing light. At the same time, the subject could be stimulated with trains of clicks. He need not pay attention to the clicks. Then the situation would be reversed. The subject would have to detect a lower-intensity click and the light flashes would be irrelevant. If selective attention works, then the bigger responses should occur to those stimuli to which the subject is paying attention. Such results were obtained by Spong and his associates in an experiment described anew by Näätänen.

Näätänen likes neither the experiment by Spong et al. nor many other experiments which appear on first reading to reflect the operation of a selective filter. The basis for his dislike is that in many cases the stimuli were presented periodically. Hence the subject could anticipate the occurrences of relevant stimuli. This anticipation or preparatory reaction could lead by itself to differential enhancement of the evoked response. Näätänen likens anticipation to a generalized arousal reaction similar to that which underlay the results of Hernández-Peón. Such a generalized reaction would not be sensory-specific and it would not be stimulus-specific. Näätänen fails to note, however, that in the Hernández-Peón experiment the evoked responses were generally diminished, while in the human work the hypothetical general process leads to a heightening of the response. He does not discuss this difference. Nevertheless, some general process could be the basis for the human results.

To remedy the faults he found in the earlier work, Näätänen did some experiments of his own. In one experiment he used procedures similar to those employed by Spong et al. The only difference was that he randomized the order of stimulus presentation, thus removing the possibility of a consistent arousal or preparatory reaction. This seems to have destroyed any selective effect: irrelevant clicks gave as large a response as relevant clicks as long as the subject was paying attention to something, even a light flash. In another experiment, Näätänen presented clicks in a regular temporal sequence and found that relevant clicks were more likely to produce a larger response than were irrelevant clicks. This procedure allowed anticipation to occur.

To support his idea that such results are due to a general effect of arousal rather than that of selective attention, Näätänen points to the fact that the enhancement effect consists largely of a change in the low-frequency components of the evoked response complex. It can be proven that any periodic function may be considered to be the sum of many simple sinusoids. It is the low-frequency components of the complex wave comprising the evoked response that contribute most of the enhancement effect. A slow negative shift in the EEG has been found to be a correlate of arousal. Also, gradual increases in cortical excitation are associated with activation by the reticular system. This leads to the suspicion that the enhancement effect may be associated with the CNV being superimposed on the evoked potential.

The experiments done by Näätänen were ingeniously designed and skillfully executed. The reader can learn something about craftsmanship from this monograph. Näätänen is a perceptive observer of his science and has pointed out potential pitfalls in the conduct of evoked potential studies. On the conceptual level, however, he is somewhat less than satisfactory. Lest these last few words be misconstrued, let me hasten to say that I admire his monograph very much. It will play an important role in the study of electrical phenomena associated with attention. My reservations stem from a deep-seated discomfort with the whole idea of using evoked potentials to study attention.

One of my reservations stems from the fact that electrical brain activity is very complicated. One does not quite know what to measure. There is no theoretical basis for saying that this or that component of the evoked potential is relevant, while some other component is not relevant. It is all very empirical at this point in history. Also, different stimuli may give very different results. Riggs and Whittle once did an experiment on binocular rivalry and failed to find an effect when the field was flashed on and off. I also did this experiment, and so did Cobb and Ettlinger. More recently, however, a number of experimenters using a different kind of visual stimulus have found a strong effect of binocular rivalry on the evoked potential. Similarly, how do we know that if different and more complex stimuli were used, an effect of selective attention would not appear? That is the trouble. There are just too few grounds for making general statements from experiments in this field.

Suppose that in some experiments an effect of selective attention were to be found. I am reasonably certain that the subject could tell us if he were more or less attentive. He might even give an estimate of the magnitude of his state of attention. Has any new information been obtained by measuring the evoked response? This reviewer is hard-pressed to answer why it is so much more impressive when a wave reflecting a voltage change correlates with a stimulus than it is when a subject behaves in some other fashion in response to the stimulus. Näätänen does not answer this question either.

One of the basic problems of science is to learn to ask the right questions—questions that will be fruitful in suggesting some deeper understanding of the phenomena being studied. Titchener once taught that attention is merely the degree of clarity of a sensation. This is really the filter theory and it was stated long before psychologists acquired engineers' terminology. Evoked responses, behavioral indexes, and recent work on the arousal systems really tell us very little more about the fine structure of the hypothetical filtering process than that there is more or less excitability or responsiveness to some elements of a situation. In most of these studies it appears that the authors assume that attention affects incoming signals—that it has its influence on the sensory side of things. Yet the band pass of the sensory system is truly astounding. Also, responses transmitted along sensory pathways are detectable in anesthetized subjects, and only the slow waves are left out. These slow waves could hardly inform us about the complexities of mental life. They do not have sufficient degrees of freedom. Perhaps we are looking in the wrong direction in this field of inquiry. Perhaps attention involves a shifting of the band of efferent activity rather than a change in input activity. This old idea seems to have barely impressed workers in evoked potentials. It is Näätänen's insensitivity to these possibilities that gives me pause. The work of Norman, of Deutsch and Deutsch, of Neisser, of Hochberg, and of Festinger and his colleagues seems to be very relevant to the problem of selective attention. Ideas concerning the location of the filter, and about the relationship between attention and perception, are conspicuous by their absence from this monograph. An enlarged conceptual framework might yet permit us to demonstrate and measure the effects of selective attention. This has not yet been ruled out by Näätänen.

LLOYD KAUFMAN, *New York University*

PSYCHOLOGICAL STRESS AND THE COPING PROCESS. By RICHARD S. LAZARUS. New York: McGraw-Hill, 1966. Pp. 466, xiii.

Some attempt was needed to bring order to the many diverse topics that belong under this title. Lazarus has given us a powerful and important try. His book is conceptually rather than methodologically oriented, and its aim is to place the topic of the psychology of stress in the broader context of the other research topics that have grown up alongside it: anxiety theory, studies oriented toward Freudian defense mechanisms; studies of extreme situations; experimental studies of therapy; experimental analyses of aggression; the whole new tradition of social psychology oriented to emotion and to social-psychological pharmacology, along with the blooming research on the physiological and biological bases for these processes and on the relations among the psychological, physiological, and even sociological processes of arousal, stress, and adjustment (or 'coping'). This book is written so that undergraduate students can enjoy it, and it is graced by inclusion of long and effective quotations from primary sources. All chapters have concise summaries, and the text is dotted with important distinctions and general principles carefully set off in italics. I have used the book for teaching with students whose first language was not English. They enjoyed the book and felt that it was clear and useful.

This enjoyable clarity may be due, in part at least, to the self-conscious human bias taken by the book. Lazarus, though not opposed to a general theory of psychological stress that will hold for all animals, does not want to "eliminate categories which are prime theoretical tools for understanding the person" (p. 15). The theme, therefore, is to broaden the conceptual net in the direction of Magda Arnold's phenomenological recommendations and to elaborate mediating processes whenever this is necessary to include the qualitative as well as quantitative differences that occur in the same person's experiences in the same place at different times or under different conditions, and in different persons' experiences in different cultures. The general stance of the book is perhaps best revealed, and clearly controversially, when the author says, in adopting the human person as the point of reference, "that the most fruitful direction from which to understand psychological stress is through its most complicated instances rather than its simplest" (p. 18). There are none of the pretensions here of the hypothetical-deductive systems, nor is there any search for the kind of explicit theory that is palpably simplistic and vulnerable.

The book first deals with the general topic of stress and its history in physiology, psychology, and other social sciences. Recognizing the confusion among levels of conceptualization of the topic (particularly the simplifications carried over into psychology from Selye's formulations), a terminology of stress is proposed and adhered to thereafter. 'Stress' becomes a collective term for an area of study rather than any part of a particular process. Psychological, physiological, and sociological levels are sharply and clearly distinguished (later elaborations probably fail to reintegrate these levels adequately, but that is a task well left to other workers). Lazarus's own status, and the impact of the book itself, may just manage to achieve some standardization of terminology in this area, and this alone would make the book worthwhile.

When a "stress stimulus" has been psychologically "appraised" by an actor, it becomes a condition of "threat": in fact, it is suggested that we use the term "threat stimulus" to refer to that which produces threat, to keep it clearly distinguished from a "noxious stimulus" at the physiological level. Activity that arises to deal (functionally) with threat is called "coping," or the "coping process(es)," and this coping intervenes between threat and the observable response. Threat becomes a major intervening process in this scheme. The discussion focuses around two aims: to persuade us that "threat appraisal," not "anxiety arousal," is the important process and to place the problem of awareness in perspective with the help of several valuable examples from the research. In dealing with the first of these aims, the criticism of anxiety theory includes a general criticism of associationistic approaches, largely on grounds of incompleteness, but also on the grounds that although Lazarus's own rich array of concepts (which distinguish between *many* emotions, defenses, attacks, avoidances, etc.) can be translated back into "traditional" terms, one is not *led* to these distinctions by associationistic thinking. In dealing with the issue of awareness, Lazarus, like Lewin, does not hold that threat appraisal need be conscious or in full awareness, and goes on to elaborate the bases of different levels of awareness.

Next, in Chapters 3 and 4, Lazarus deals with the stimulus factors and personality factors that determine the appraisal of threat. The discussion of "stimulus factors" deals with some pretty complex matters, such as the ratio of the threat (as appraised) to the means for meeting the threat, and the conditions under which ambiguity is adequate or inadequate to permit threat to be appraised. The discussion of "personality factors" includes motivational characteristics, belief systems, and the intellectual resources of the individual appraiser. This chapter is particularly restrained, self-critical, and suggestive.

So far, the conditions causing the primary appraisal of threat have been exemplified. Discussed next is the "process of secondary appraisal," which underlies the selection of a coping strategy. This particular form of cognitive activity intervenes between threat and coping, to determine whether, for example, anger, fear, direct action, or defensive processes occur. The important aspects of secondary appraisal dealt with in Chapter 5 include the degree of threat and the factors in the stimulus configuration that determine *which* coping reaction occurs. This discussion, with relevant research examples, is followed by a discussion of the personality factors that are involved in secondary appraisal and coping choice, ranging from motivation patterns, through ego resources and one's "coping dispositions," and emerging with general beliefs about the environment and one's available resources.

Chapter 7 shifts to a more systematic-classificatory mode, and the author makes a very interesting contribution in classifying coping-reaction patterns and the appraisals that underlie them. The analysis focuses on four main kinds of coping patterns: (*a*) actions aimed at strengthening the individual's resources against harm, (*b*) attack, (*c*) avoidance, and (*d*) defense, with each of these being broken down into a number of subvarieties with their own forms of appraisal. Perhaps these pages provide the most persuasive case for the value of an otherwise fairly rococo conceptual scheme.

Chapters 8 and 9 turn to methodological matters involved in the measurement of psychological stress, covering reported affects, motor-behavioral reactions, and physiological reactions. These chapters provide middle-range students (and people with other specialties) a good glimpse at the problem and the generally slippery quality of measurement in this area, and they include a good introduction to the general problem of patterning in physiological measures.

The book ends in a strongly integrative mode. Chapter 10 includes a neat flow-chart arrangement, which (consistent with the book's general aims) outlines the flow of psychological stress and the flow of physiological stress (borrowing from Selye for this latter, of course). The parallels between these two general processes are outlined, while they are kept completely separate analytically. The major point of convergence between the physiological and psychological systems is pointed out: there exist common and practically identical physiological responses to both psychological and physical antecedents. Further, if they are perceivable, these physiological responses act as stimuli to be psychologically appraised, and so on. The bias is psychological, however. Physiological stress pressures are not, for example, given equal status as having independent direct effects upon psychological systems. The perspective, however, is that of behavioral science, and this would have been more clear if a flow chart for sociological stress had been included.

If anything, this book overshoots conceptually, by positing more entities, mediators, and qualitatively differentiated processes than are needed, but in so doing it achieves its aim of integrating a messy field and makes explicit a model for an area where models are often implicit. We are left with a sense that there is an honest fit between this phenomenologically oriented model and a field which still operates on the basis of a commonsense model. Still, Lazarus's argument is based on a moot point, I think. He feels that students and fellow researchers are motivated to explore when the conceptual scheme is rich and close to immediate experience. This is probably true for those of us who are 'integrators.' But there are other tastes, and it would be surprising if a good many readers were not stimulated to get to work in another direction, to show why a good deal of the conceptual machinery massed here is unnecessary because simpler and more parsimoniously elegant models can work as well or better—and lead us on to at least as much fresh discovery—as the present theory.

William W. Lambert, *Cornell University*

SENTENCE STRUCTURE AND THE READING PROCESS. By I. M. Schlesinger. The Hague: Mouton, 1968. Pp. 171. $5.75.

I. M. Schlesinger describes his doctoral research against a background of other studies, all related to the effect of grammatical structure on reading. Although it seems intuitively plausible that some sentences are harder to understand than others because of their different grammatical form, most reading research has concentrated on lower levels of structure as sounds, letters, and words. In the design of elementary readers, there is usually careful control of vocabulary, and sometimes of phoneme-grapheme relationships, but grammatical structures are almost randomly selected, whether for complexity or for frequency in children's speech.

One reason for the neglect of grammar as a variable has been that the science of linguistics has only recently begun to supply a systematic basis for understanding grammatical differences or for predicting their effects. Length of sentences and number of subordinate clauses were almost the only grammatically relevant variables that were both measurable and plausibly related to reading rate, and both of these are confounded with so many other grammatical dimensions that it is impossible to separate out the specific causes of comprehension difficulty. Grammatical theories, such as transformational ones and those developed in connection with computer translation, have provided a basis for much more specific, and therefore more interesting, studies of linguistic comprehension, including reading.

Schlesinger works with the general hypothesis that linguistic theories correspond to the psychological processes involved in using language and that, therefore, units such as phrase structure and linguistic dimensions of complexity and depth have their counterparts in psychological variables. He reviews research by himself and others on a number of specific hypotheses, most of them derived directly from linguistics. One set of studies uses measures of eye-voice span (length of material stored in memory between eye fixation and voicing in oral reading) to show that units in memory storage correspond to phrase-structure units. Other studies show that making transformations (most of those studied being passive, negative, and question transformations) increases the difficulty of understanding as measured in a number of different ways. Although these transformation studies were not satisfactorily controlled for sentence length, a further study showed that sentence length had no effect when it was isolated from other variables. It had been predicted that sentences in which clauses are nested so that several parts must be stored in memory before they can be linked with related constituents would be more difficult to comprehend because of the strain on memory. In fact, there was only minimal evidence of such an effect, and then only when the subject was prevented from reviewing earlier parts of the sentence.

Up to this point Schlesinger has been proceeding on an assumption implicit in much of linguistic theory—that syntactic structure should and could be examined independently of semantic considerations. But carrying the linguistic assumption over into psychological questions suggests that syntactic variables have their own effects on language processes. The results of the seemingly well-constructed experiments with nested sentences lead to a change in that assumption and to a new hypothesis—that semantic decoding processes interact with syntactic ones. Some clever experiments with nested sentences with and without semantic cues to decoding confirm the intuition that syntactic factors are not primary nor independent. (And in the years since these studies were done, linguistic theories have developed in the same direction, so as to take better account of the conditions that Schlesinger designates as semantic factors.)

The traditional ideal of pure research in which hypotheses were based entirely on theory has limitations that are well illustrated in this research report. The research hypotheses were theoretically sound and the experiments appropriately designed and controlled. The results, however, were divided between those that gave partial, limited support for the hypotheses and those that pointed in other directions and required basic revisions of the underlying theoretical assumptions. One of the reasons why these other possibilities were observed and tested was that investigators were concerned with the practical problems of readability and reading skill as well as with theory. The study of reading cannot be divorced from the study of meaning and still have any practical value. Thus while sound theoretical considerations suggest avoiding semantic factors and concentrating on behaviors that could be studied independently of them, an experimenter whose motivation includes the desire to be 'relevant' is inevitably led to measures and conclusions that depart from theoretical purity but ultimately bring the theory itself into better correspondence with reality. At the same time, Schlesinger rightly argues, the broader conceptual orientation in linguistics and communicability has made a valuable contribution to readability research, a contribution that a more purely practical approach could not have made.

JANE W. TORREY, *Connecticut College*

SYMBOLS AND SOCIAL THEORY. By HUGH DALZIER DUNCAN. New York: Oxford University Press, 1969. Pp. 314. $7.95.

In this book Dr. Duncan has made an effort to provide a historical and theoretical background for symbol-interaction theory. This is a theory that strives to understand order (and disorder) and continuity (and change) in social processes by means of an analysis of social communication—with primary emphasis on the symbol as the basic unit of analysis. The symbol is seen as that which determines and structures social relations; that which defines roles and positions in society; that which forms the bridge between the form and content of social action. In this book, however, all of these statements remain programmatic, general and incomplete. Nowhere is there a clear and coherent statement of how political, social, religious, and economic symbols actually emerge and operate on, or in, social institutions or groups. This leads to a very unsatisfactory state of affairs for the reader who is not entirely familiar, or in agreement, with the theory. There can be no doubt that the issues raised in the book are crucial (i.e., the relations between form and content, impression and expression, order and change), but there is no real sense of resolution or progress in knowledge. One may again affirm the importance of the problems without in any way arriving at a new point of departure or at a new perspective from which to gain insights and clarifications. Nor is there any doubt that the theoreticians discussed in the book are, by any standard, of first-rate importance (i.e., Weber, Pareto, Durkheim, Tonnies, etc.), but they are necessarily treated in a very selective fashion. They remain skeletons and specters, not only within their own theoretical frameworks but also within the context of their contributions and connections to symbolic interaction theory.

Those readers who are already familiar with the social-symbolic school of thought (either from the American side of G. H. Mead and Kenneth Burke, or from the European side of Max Weber and Emile Durkheim) may well find that this book makes an important addition to their understanding in that it exposes common roots and makes heretofore-unknown relationships clear. For the large group of undecided or unconvinced readers, however, this book is likely to be one whose initial appeal (on the grounds that the issues raised are real and the people discussed important) leads to ultimate disappointment.

AARON HERSHKOWITZ, *Yeshiva University*

BEHAVIORAL PHARMACOLOGY. By TRAVIS THOMPSON and CHARLES R. SCHUSTER. Englewood Cliffs, N.J.: Prentice-Hall, 1968. Pp. ix, 297.

This book can be divided into three sections. The first is a useful primer into some of the basic concepts of pharmacology: dose-response curves, how drugs are absorbed and distributed in the living organisms, and abbreviated reviews of the basic pharmacology of many of the drugs that alter behavior. The second section reviews methods of studying animal behavior, with heavy emphasis on operant behavior. The final section is concerned, for the most part, with the action of various centrally acting drugs on operant behavior. Primarily because of the research interests of the authors, there is a very good subsection on the self-administration of drugs by animals where the reinforcer is the drug itself. The appendixes are useful, for they detail various pharmacological techniques that the psychology student would find difficult to obtain through other sources. Their descriptions are accompanied by a number of good photographs.

The dust cover states that the book "examines principles and techniques of experimental psychology and pharmacology and applications to behavioral actions of drugs." If one defines experimental psychology as limited to operant behavior, then the book meets its goals. However, there are behavioral studies with drugs and useful experimental methods for the study of effects of drugs on behavior that are either ignored or given very little emphasis in this volume.

CONAN KORNETSKY, *Boston University School of Medicine*

ANNOUNCEMENTS

AMERICAN PSYCHOLOGICAL ASSOCIATION MEETING FOR 1971

The American Psychological Association Board of Convention Affairs announces the "Call for Programs" for the 1971 Annual Convention. Because of the ever expanding interests of the membership, the number of hours requested by the divisions for substantive programs has increased almost 18%, thus necessitating complex multiple tracking. Since there are many areas of overlap between the divisions, the Board urges the divisions to cooperate in planning their programs. Sequential programming time may be requested by divisions if they feel it would enhance development of a theme or subject area.

All correspondence concerning divisional programs must be channeled through the appropriate division program chairman, to whom all proposals must be submitted by February 15. Contact the Board of Convention Affairs for further details.

SENIOR FULBRIGHT-HAYS PROGRAM FOR 1972–1973

The Committee on International Exchange of Persons announces that applications for senior Fulbright-Hays awards for lecturing and research during 1972–1973 in over 75 foreign countries will be accepted in the spring of 1971. Specialists in the field of psychology who are U. S. citizens and have a doctorate or college teaching experience are invited to indicate their interest in 1972–1973 awards by immediately completing a simple registration form, available on request from:

Senior Fulbright-Hays Program
2101 Constitution Avenue
Washington, D. C. 20418

Registrants will receive the detailed announcement of available awards as soon as it is issued in the spring, in time to consider the possibilities and to apply before the closing date.

Senior Fulbright-Hays awards generally consist of a maintenance allowance in local currency to cover normal living costs of the grantee and family while in residence abroad, and round-trip travel for the grantee (transportation is not provided for dependents). For lecturers going to most non-European countries, the award includes a dollar supplement, subject to the availability of funds, or carries a stipend in dollars and foreign currency, the amount depending on the assignment, the lecturer's qualifications, salary, and other factors.

INDEX

RAND B. EVANS, Wright State University

AUTHORS

(The names of authors of original articles are printed in CAPITALS and SMALL CAPITALS; of authors of books reviewed, in roman; and of reviewers, in *italics*.)

Adams, J. A. 299
Alker, H. 457
AMSTER, H. 22
ANDERSON, N. H. 567
ATTNEAVE, F. 1
Ausubel, D. P. 303

Bancroft, T. A. 140
Bauer, R. A. 458
BAUST, R. F. 130
Beck, J. 448
Beecroft, R. S. 293
BERGER, S. M. 86
BINTZ, J. 283
BIRNBAUM, I. M. ...412
Bock, R. D. 289
Borden, G. A. 454
Brier, R. 304
BRISLIN, R. W. 372
Brown, D. R. 292
Bugelski, B. R. 301
Burroughs, W. S. ...298
Bush, R. R. 296
Butcher, H. J. 455

Carterette, E. 448
Chernoff, A. 142
CHITAYAT, D. 352
CHRIST, R. E. 264
Church, J. 294
COLEGATE, R. L. 208
Cook, H. 303
COREN, S. 189
Coren, S. 291
Corso, J. F. 445
Cotman, C. 452
Craig, J. C. 145
CULVER, C. M. 420
CURTIS, C. D. 256
CVETKOVICH, G. 119

DABICE, M. 613
Davitz, J. R. 454
DEY, M. K. 248
Driver, M. J. 136
DUNCAN, C. P. 535
Duncan, H. D. 639

Eden, M. 292
EPSTEIN, W. 322
ERBER, J. T. 428
ERIKSEN, C. W.
208, 330
Evans, W. O. 296

Fairweather, G. W. 302
FESTINGER, L. 189
FISHBEIN, H. D. 544
FREUND, J. S. 343
FRIESON, D. 544
FROMMER, G. P. 86

Galanter, E. 139
GARDNER, G. T. 40
GARNER, W. R. 151
Gergen, K. J. 458
Gilman, A. 298
GOGEL, W. C. 55
Goodman, L. S. 298
GOTTWALD, R. L. ...151
Gregg, R. B. 454
Gross, C. G. 453
Grove, T. G. 454
Gurowitz, E. M. ...452

Haber, R. N. 138
HALL, J. W. 272
HALL, J. F. 436
Harris, D. B. 455
Hay, J. C. 138
HAYGOOD, R. C. 544
Hellmuth, J. 142
Hershkowitz, A. 639
HINTZ, J. M. 126

Hoffer, A. 298
Holding, D. H. 280
Hood, D. C. 453
Houston, J. 299
Hyman, R. 136

Irwin, D. S. 86

Jones, L. V. 289
Jones, M. B. 222

Kanak, N. J. 256
Kappauf, W. E. 630
Katahn, M. 237
Kaufman, L. 634
Kausler, D. H. 428
Kaye, H. 68
Kempler, B. 68
Kenshalo, D. R. 145
Keppel, G. 22
Kerr, T. P. 112
Kline, N. 296
Köhler, W. 291
Kolers, P. A. 292
Kornetsky, C. 640
Kozloff, E. E. 272
Krantz, D. 289
Krueger, L. E. 389

Lambert, W. W. 637
Lazarus, R. S. 634
Leibowitz, H. W. ..372
Lester, G. 554
Lindauer, M. S. 130
Lingeman, R. R. ...298
Lunzer, E. A. 294

McGee, V. E. 140
McGuigan, F. J. ...377
Mack, A. 352
Mackay, S. A. 311
Malpass, R. S. 302

Martin, E. 311
Marx, M. H. 452
Massaro, D. W. 567
Masters, L. 76
Masters, R. E. L. ..299
Mays, L. 436
Medical Economics,
Inc. 299
Mershon, D. H. ... 55
Mewhort, D. J. K. 439
Meyer, A. 22
Michalewski, H. ..283
Miller, A. L. 95
Moles, A. 451
Morant, R. B. 554
Morgan, C. L. 322
Morris, J. F. 294
Mueller, J. H. 401

Näätänen, R. 631
Nelson, T. M. 126

Olson, G. A. 428
Olson, R. K. 1
Osmond, H. 298

Parker, J. F. 171
Parks, T. E. 442
Pedersen, V. M. .. 95
Pickering, J. P. ...401
Plutchik, R. 301
Postman, L. 171
Preusser, D. 151

Rachlin, H. 452
Rappoport, L. 119
Raskin, D. C. 293
Ravanette, A. T. ..142
Rhine, J. B. 304
Rivers, G. 189
Rohrbaugh, J. W. .330
Rosevear, J. 299

Schlesinger, I. M. ..637
Schroder, H. M. ...136
Schuster, C. R. 639
Scott, K. G. 142
Sheldon, R. W. 95
Simon, H. A. 139
Singer, R. 141
Smith, M. B. 457
Snyder, C. R. 237
Solomon, D. 298
Solso, R. L. 216
Sperling, G. 461
Sem-Jacobsen, C. W.
143
Stafford, J. W. 286
Streufert, S. 136
Sullivan, R. 613
Summers, D. A. 304
Szasz, T. S. 141

Teghtsoonian, M. .601
Teghtsoonian, R. ..601
Thompson, T. 639
Thurlow, W. R. ..112
Torrey, J. W. 639
Turnage, T. W. ...621

Underwood, B. J. ..343

Valle, F. P. 103
Vitz, P. 451

Warren, R. 613
Warren, R. M. 448
Warren, R. P. 448
Weick, K. 458
Weintraub, D. J. .. 40
Welford, A. T. 299
Wenderoth, P. M. .576
Williams, J. 142
Wilson, W. A., Jr. ..143

Zeigler, H. P. 453

SUBJECTS

(References in *italic* figures are to reviews.)

Accommodation, binocular: physical model for, 477–482; fusion and, physical model for interaction between, 492; neural mechanisms for, 462, 468, 496, 527–529; vergence and, in physical model for interaction between, 487–491

Acoustic confusability, effect on visual search, 389–400

Adaptation: perceptual; eye-dependent, 352–371; visual, to tilt; interocular transfer of, 354

Adaptation level, noise-aversion thresholds and, 614–620

Aesthetic perception, information theory and, *451 f.*

Affect, complex verbal learning and, 237–247

Aftereffect, surface-slant: contrast effects and, 577; dissipation of, 580 ff.; fixation time and, 577–580; monocular vs. binocular observation and, 585 f., 589; stimulus contours and, 583–586, 589 f.; surface texture and, 582–586, 589 f.; visual, size of, Emmert's law and, 40–47, 51–54

Age, size constancy and, 372–376

Animal: Rat, agouti-selfed; defecation rate in, 103 f., 108 f.; locomotion rate in, 103–106, 108–111; open-field behavior of, 103–106, 109 ff.; rearing frequency of, 103–106, 109 ff.; sex differences in, 103–111; thigmotaxis in, 103 f., 107, 109 ff.; black-hooded; defecation rate in, 103 f., 108 f.; locomotion rate in, 103–106, 108–111; open-field behavior of, 103–106, 109 ff.; rearing frequency of, 103–106, 109 ff.; sex differences in, 103–111, thigmotaxis in, 103–105, 107, 109 ff.; strain of; open-field behavior and, 103–106, 109 ff.

Anxiety, state: complex verbal learning and, 237–247

Apparatus: brightness-contrast, 58; eyeblink transducer, photoconductive for humans, 283 ff.; size-distance judgments and Emmert's principle of perceived size, 48; variable-interval, 283

Artificial phenomena, *139 f.*

Associations: intertriad, feedback and, 401–411; number of, problem solving and, 535–543; position, generalization of, in rote serial learning, 248–255; strength of, in paired-associate learning, 22–39; context-item, unlearning of, 256–263; first-list, maintenance of, during transfer, 171–188; general vs. technical, temporal changes in, 627 f.; specific, unlearning of, 256–263. *See also* Learning: Paired-associate

Association measures, five-letter words, 436 ff.; interrelations among, 436 ff.

Attention: brightness assimilation and, 189–207; brightness contrast and, 189–207; saliency and, in problem solving, 545 f.; selective, evoked potentials and, *631–634;* selective, factors influencing, 330–342; selective interstimulus interval and, 332–339; spatial separation between elements and, 339–342

Aversion, noise: thresholds for, adaptation level and, 614–620; in humans, 613–620

Binocular combinations: Hochberg stereograms and, 519–521; Kaufman figures and, 517 ff.; stabilized vision and, 521 ff.
Binocular rivalry, neural model of, 502–505, 508–511, 515 ff.
Biography: Pace, E. A., 286; Wundt, W., 286. *See also* Necrology
Brain stimulation, in humans, *143 ff.*

Choice, measurement of, 289 ff.
Cognitive differences, in mixed-motive conflict situations, 119–125
Communication, speech behavior and, *454 f.*
Computers, digital, *139 f.*
Concept learning, relevant and irrelevant saliency in, 544, 553
Conditioning, classical, *293 f.*
Conflict situation, mixed-motive: cognitive differences in, 119–125; reward structure and, 119–125
Constancies: brightness; Gelb effect and, 55 ff., 60, 63 ff.; instructions and, 130–135; binocular cues and, 55–67; depth cues and, 62–67; size; size and distance effects on, 601–612
Contour, perceptual grouping and, 2
Control, experimental, 139 f.; in paired-associate learning, 24–27; in perceptual grouping, 2 f., 6, 10; in social innovation, *302 f.;* ocular tracking, in postretinal visual storage, 442 ff.
Cortical brain fields, Gestalt view of, *292.*
Cortical dominance, visual form recognition and, 593 f.
Cortical locus, of visual form recognition, 591–594, 599

Cues, depth: on whiteness-contrast judgments, 62–67; stereoscopic, on whiteness-contrast judgments, 55–67

Defecation rate, strain differences in rat, 103 f., 108 f.
Disparity-depth relationship, modification of: with uniocular image magnification, 322–329
Drugs: amphetamine, effect on cognitive skills in humans, *296;* behavioral effects of, *639 f.;* chlorpromazine, effect on cognitive on skills in humans, *296;* hallucinogenic, 296–299; LSD, effect on humans, *296 f.;* marihuana, *298 f.;* meprobamate, effect on humans, *296;* psychoactive, *296 f.;* therapeutic, pharmacological basis of, *298;* tranquilizers, *296 f.*
Dynamic psychology, T. V. Moore and, 287

Educational psychology, cognitive view of, *303 f.*
Electromyographic activity, observational learning and, 86–94
Emmert's law of retinal size, size-distance judgments and, 40–54
Emmert's principle of perceived size: apparatus for testing, 48; size-distance judgments and, 40–54
Errata, 285, 444
Estimation. *See* Judgment
Evoked potentials, selective attention and, *631–634*
Eyeblink transducer, photoconductive, for humans, 283 ff.
Eye movements, visual form recognition and, 596–599
Eyestrain, mediation by, in visual form recognition, 596–599

Feedback: effect in verbal-discrimination tasks, 401–411; complex verbal learning and, 237–247; in

motor-skill acquisition, 222–236
Figures, ambiguous: locus of recognition in, 591–600
Figure-ground; in brightness assimilation, 190 f.; in temporal patterns, 163–168
Forgetting, interference theory of, 412 ff., 418 f. *See also* Retention
Fusion: accommodation and, physical model of interaction between, 492; binocular, mechanism for, 462–465; horopter and, 482; neural model of, 506 f.; physical model for, 482

Gelb effect, brightness constancy and, 55 ff., 60, 63 ff.
Gestalt psychology, *291 f.*
Golden section, perimetric hypothesis and, 126–129
Guessing: errors in, response-strength distribution and, 535–543; order-of-approximation effect and, 439–442

Handwriting: as linguistic behavior, 377–388; breathing rate during, 380 ff.; quality of, covert oral behavior and, 377, 381–388
Head tilt, perceptual grouping and, 7–21
Helmholtz: on ocular divergence, 463; on perception, *448–451*
Horopter, binocular fusion and, 482 f.

Illusion: Audiogyral; differential-hearing-change hypothesis of, 555, 562, 565 f.; frame-of-reference hypothesis of, 555, 562–566; perceived head position and, 556, 562–566
  Line; depth cues and, 280 ff.
  Geometrical; perspective theory of, 567–575
  Müller-Lyer; criticism of Gregory's theory of, 280 ff., 567–575
  Ponzo; size constancy and, 372 f.
Information processing, human, *136 ff.*
Information theory, aesthetic perception and, *451 f.*
Intelligence, human, *455 ff.*
Interference, linguistic: effect on retention and learning, 22–39

Judgment: Depth; whiteness contrast and, 62–67
  Grouping. *See* Perceptual grouping
  Length; average, 95–102;
  Magnitude-of-size; college vs. high-school subjects in, 602–611; size and distance effects on, 601–612
  Size-distance; Emmert's law of retinal size and, 40–54; Emmert's principle of perceived size and, 40–54
  Whiteness-contrast; depth cues and, 62–67; stereoscopic cues and, 55–67

Labeling: theory of, 73 ff.; trigram, for random shapes, 68–75
Language tasks, covert oral behavior and, 377–388
Learning: Concept; relevant and irrelevant saliency in, 544–553
  Development of; in humans, *294 f.*
  Disorders of, *142 f.*
  Materials; five-letter words, association measures of, 436 *ff.*
  Motor; acquisition of, feedback and, 222–236; reproductive, electromyographic activity in, 86–94
  Observational; electromyographic activity in, 86–94
  Paired-associate; association strength and, 22–39; experimental control in, 24–27; context-item, 256–263; interference in, children vs. adults, 22–39; interpolated reading and, 412–419; item associability and, 76–85; knowledge of

results and, 76–85; letter-sequence interference in, 23, 28 ff.; linguistic interference in, 22, 28 ff.; linguistic environment and, 621–629; list-differentiation hypothesis in, 311–321; natural language mediators and, 77–85; 'natural' situations and, 621–629; recall clustering in, 317–321; recall frequencies in, 315 f.; recall ordering in, 316 f.; retroactive inhibition, long-term, in, 621–629; specific, 256–263; stimulus selection in, by children, 216–221; transfer effects in, 626–629; transfer effects, negative, 311–321
  Processes of, *452*
  Recognition, false; number of presentations and, 272–279
  Rote serial; generalization of position association, 248–255
  Serial; generalization of position association, 248–255; intrusions and, 253 ff.
  Transfer of, 311–321; first-list associations and, 171–188; habit substitution vs. habit accretion in, 171 ff., 181–188; paired-associate, 311–321; paradigms for, compared, 171–188; simultaneous vs. successive methods in, 171–188
  Verbal; affect and, 237–247; feedback and, 237–247; state anxiety and, 237–247
  Verbal-discrimination; intertrial associations and, 401–411; feedback and, 401–411; implicit response units and, 401–411; selection strategies in, 428–435; taxonomic categories as right or wrong items in, 428–435; mediation in, 428–435
Linguistic behavior, handwriting as, 377–388
Localization: auditory; in moving visual field, 112–118; vestibular stimulation and, 554–566; tactile; eye dominance and, 421 f., 425; hand dominance and, 421 f., 424; inhibitory processes and, 426 f.; perspective hypothesis of, 426 f.; visual double-image suppression and, 427; visual form recognition and, 591–600
Locomotion rate, strain differences in rat, 103–106, 108–111

Meeting: Society of Experimental Psychologists, 630
Memory, *139 f.;* molecular basis of, *452 f.;* recognition, short-term, word frequency and, 343–351; RNA and, *452 f.;* short-term, *301. See also* Retention
Method, experimental designs and, *140*

Naming: of random shapes, with trigrams, 68–75; similarity of perception and, 74 f.; theory of, 73 ff.
Necrology: Thomas Verner Moore, 286
Neuron-pair interactions in binocular vision, 529–534
Neurophysiology, *453 f.*
Noise, aversion thresholds for: in humans, 613–620; adaptation level and, 614–620
Nonsense trigrams as labels for random shapes, 68–75

Ocular divergence, Helmholz on, 463
Ocular tracking, control for: in postretinal visual storage, 442 ff.
Open-field behavior in rat, 103–111
Order-of-approximation effect, guessing and, 439–442

Panum's fusional area, 486 f.; binocular vision and, 486 f.
Parapsychology, *304 f.*
Perception: Aesthetic; information theory and, *451 f.*
  Auditory; localization, in mov-

ing visual field, 112-118; vestibular stimulation and, 554-566
Binocular; depth, evolution of, 523 ff.; golden section and, 126–129
Brightness; assimilation, adaptation level and, 193; in grouping, 2; instructions and, 130–135; practice effects and, 130–135
Brightness-assimilation; figure-ground effects on, 190 f., 193–196; method of measurement and, 196–205
Brightness-contrast; effect of attention on, 189–207; method of measurement and, 196–205
Constancies; brightness, binocular cues and, 55–67; depth cues and, 62, 64–67; size, size and distance effects on, 601-612
Depth; disparity and, in uniocular image magnification, 322–329
Figure; ambiguous, recognition and, 591–600; and ground, in temporal patterns, 163–168; in perceptual grouping, 4
Form; encoding of, by implicit speech, 208–215
Helmholtz on, 448–451
Law of apparent area, 602
Line-length; depth cues and, 280 ff.
Organization of, in temporal patterns; figure and ground in, 163–168; one-element, 161–161, 167–170; gaps and, 157–160; run order and, 151–161, 167–170; two-element, 151–156, 161–170
Pattern; in men and machines, *292 f.*
Size; magnitude estimation of, size and distance effects on, 601–612
Size-constancy; age and, 372–376; object separation and, 372–376; Ponzo illusion and, 372 f.
Size-distance; Emmert's law of retinal size and, 40–54; Emmert's principle of perceived size and, 40–54
Tactile; localization and, errors in, 420–427; eye dominance and, 421 f., 425; hand dominance and, 421 f., 424 f.; perspective hypothesis of, 426; visual double-image suppression and, 427
Visual. *See* Vision *and* Visual
Perceptual grouping: angles, size of display and, 7, 14; brightness as factor in, 2; contour as factor in, 2; control, experimental, in, 2–4, 6, 10; curvilinear stimuli, size of display and, 7, 9, 14; figure and ground in, 4; gravitational vs. retinal orientation in, 17–21; line linearity as factor in, 3; line slope as factor in, 2, 4, 6, 9–17, 20 f.; of lines, size of display and, 7, 14; rate of, effect of stimulus factors on, 1–21; retinal vs. gravitational factors in, 17–21; shape as a factor in, 2; similarity and, 1–21; slope analyzers in, 11 f., 21; stimulus configurations and, 1–21; stimulus-field orientation and, with head tilt, 7–21
Pharmacology, behavioral, *639 f.*
Policy formation, *458 ff.*
Preferences, visual, of rectangles: golden section and, 126–129
Problem solving: number of associates and, 535–543; response strength and guessing errors in, 535–543
Psychiatry, the law and, *141 f.*
Psychopharmacology of drugs, *296–299*

Reaction time: as measure of perceptual grouping rate, 1–21;

choice, stimulus-exposure time and, 264–271

Reading: failures in, *142;* sentence structure and, *637 ff.*

Rearing frequency, strain differences in rat, 103–106, 108–111

Recognition: false, number of presentations and, 272–279; figures, ambiguous, 591–600; model for, 277 ff.; retention interval and, 347 ff.; visual-form; cortical dominance and, 593 f.; cortical locus of, 591–594, 599; eye movements and, 596–599; eyestrain as mediating stimulus in, 596–599; fixation point and, 594 ff.; 'phenomenal' vs. retinal locus of, 594 ff.; trace localization and, 591–600

Retention: associations; context items in, 256–263; specific, 256–263; interference, linguistic, and, in paired associates, 22–27, 29–39; paired-associates; children vs. adults, 22–39; effect of interpolated reading on, 412–419; recall in, immediate vs. delayed, 314; interference and, in 'natural' situations, 621–629

Retinal disparity, altered adaptation to, 352–371

Reward structure, in interpersonal conflict situations, 119–125

RNA, memory and, 452 f.

Saliency, attention and: in problem-solving, 545 f.

Scaling: of average length, 95–102; of average length, standard availability and, 99–102; psychological, *289 ff.*

Sensory behavior, *445–448*

Sensory processes, *453 f.*

Sentence structure, reading process and, *637 ff.*

Sex differences in rat, 103–111

Shape, perceptual grouping and, 2

Skill: fundamentals of, *299 ff.;* motor, acquisition and feedback, 222–236

Skin senses, *145*

Slope analyzers, perceptual grouping and, 11 f., 21

Social communication, *639*

Social innovation, experimental methods in, *302 f.*

Social psychology, human values and, *457 f.*

Social theory, *639*

Sound: apparent displacement during vestibular stimulation, 554–566; localization in moving visual field, 112–118

Speech: behavior, communication and, *454 f.;* implicit, as encoder of visual perception, 208–215

Stabilized vision, binocular combinations and, 521 *ff.*

Statistics, 140; curve fitting in, *289*

Stimulus selection in paired-associate learning by children, 216–221

Stimulus size, effect on perceptual grouping, 7, 9, 14

Strategies, selection: in verbal-discrimination learning, 428–435

Stress, psychological: adjustment and, *634–637*

Symbol-interaction theory, *639*

Tapetenbilden, interference in slant aftereffects, 589

Theory: associative-probability, of reference, 73 ff; representative, of reference, 74 f.

Thigmotaxis in rat, 103–105, 107, 108–111

Trace localization, visual form recognition and, 591–600

Values, affective, *289*

Verbal behavior, temporal shifts in: in introductory psychology students, 621–629

Vergence: accommodation and, physical model of interaction between, 487–491; fusion and,

physical model of interaction between, 491 f.; ocular, mechanism for, 462, 467; physical model for, 467–477
Vestibular stimulation, auditory localization during, 554-566
Vision: Binocular; accommodation, fusion, and physical model of interaction between, 492; accommodation, neural mechanism for, 462, 468, 496, 527 ff.; accommodation, physical model for, 477–482; combination, 511–523; depth discrimination, evolution of, 523 ff.; divergence, 463 *f*.; fusion and, horopter and, 482 f.; fusion, mechanism for, 462–465; fusion, neural model, predictions of, in, 506 f.; fusion, physical model for, 482; neural model for, 496; Panum's fusional area, 486 f.; physical model for, 465–496; theory of, 461–534; torsional rotation, physical model of, 477; vergence, accommodation and, physical model of interrelations between, 487–491; vergence, fusion and, physical model of, 467–477, 491 f.; vergence, mechanism in, 462, 467; vergence, neural model of, 507 f.
  Retinal orientation; perceptual grouping and, 17–21
  Stereopsis; Keplerian theory of, 498
  Stereoscopic; as cue for brightness contrast, 55–67
  Theory and research in, *138 f.*
  Transmission of steady-state information in, 191 ff.
Visual aftereffects of surface slant, 576–590; contrast effects and, 577; extinction of, 580 ff.; fixation time and, 577–580; monocular vs. binocular observation and, 585 f., 589; stimulus-contours and, 583–586, 589 f.; surface texture and, 582–586, 589 f.; *Tapetenbilden* phenomenon and, 589
Visual averaging, Müller-Lyer illusion and, 280 ff.
Visual disparity, monocular cues and: in disparity-depth relationship, 322–329
Visual displays, encoding of: acoustic confusability and, 389–400
Visual field: binocular, preference for rectangles and, 126–129; movement of, sound localization and, 112–118; rotation of, eye-dependent and disparity adaptation to, 352–371
Visual form, recognition of: cortical dominance and, 593 f.; cortical locus of, 591–594, 599; eye movements and, 596–599; eyestrain as mediating stimulus, 596–599; fixation point and, 594 ff.; retinal vs. 'phenomenal' locus of, 594 ff.; trace localization and, 591–600
Visual search, acoustic confusability and, 389–400
Visual storage, postretinal: control for ocular tracking in, 442 ff.

BUREAU OF EDNL. & PSYL. RESEARCH
LIBRARY